The impact of immunosuppression on nucleotide sequence diversity in the first hypervariable region (HVR1) of hepatitis C virus (HCV).

Thesis

UNRESTRICTED

THE IMPACT OF IMMUNOSUPPRESSION ON NUCLEOTIDE SEQUENCE DIVERSITY IN THE FIRST HYPERVARIABLE REGION (HVR1) OF HEPATITIS C VIRUS (HCV)

A dissertation submitted to the Open University by

Zarah Yetunde Lawal

in candidature for the degree of Doctor of Philosophy

Addenbrooke's NHS Trust

(Sponsoring Establishment)

for
National Institute ~~of~~ Biological Standards & Control
East Anglian Blood Centre
University of Cambridge Division of Transfusion Medicine

(Collaborating Establishments)

May 1996

Date of submission: May 1996
Date of award: 21 May 1997

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ERRATA

p. 101 Table 4.1 title should read "6", and not "5", OLT candidates.

p. 103 First paragraph (beginning on p. 102) should read:- "In preliminary experiments to assess the feasibility of the proposed experimental protocols, five HCV-seropositive samples (V, W, X, Y and Z) were analysed for the presence of HCV NCR or *E1/E2* sequences. Samples were also obtained from 4 apparently healthy blood donors (Ad, Wad, Fra and Dix) who had a past history of intravenous drug abuse (IVDA), and provided the data for non-transplanted controls in the main experiments."

p. 108 2nd paragraph should begin "The mean age of the non-transplanted patients above was 32±6 ..."

p. 109 Line 3 down, "Z1" should read "Z."

p. 110 Paragraph 1, lines 4 - 5: The 5 isolates quoted should be C1, L1, L2, M1 and Fra (and not Har, Wad, Ad, Z2 and W1 as written).

p. 113 Addendum to first sentence: "No HVR1 sequences could be obtained from W1 or Z)."

p. 136 In section 5.5, standard deviations for mean numbers of genomic variants per sample in immunosuppressed and untreated individuals were 1.46 and 3.06, respectively.

Line 8 of section 5.5: Sentence ending "... HCV hosts." should be followed by the phrase "These differences were not significant."

p. 151 Last paragraph should begin "*Most*", and not "all," criteria.

Table of Contents

DECLARATION

The work described in this dissertation was carried out under the direction of Professor Jean-Pierre Allain in the Division of Transfusion Medicine, Department of Haematology, University of Cambridge, and the East Anglian Blood Centre, between October 1992 and September 1995. It was carried out entirely by me, except for those parts which are explicitly indicated in the text as being the result of collaborations. This dissertation has not been submitted, either wholly or in part, for a degree, diploma or other qualification at any other university.

Zarah Y. Lawal

May 1996

Acknowledgments

First of all, I thank Professor Jean-Pierre Allain, without whom this project would never have got off the ground. During his direction of these studies, he demonstrated an exemplary sense of focus, always demanding dedication, accuracy and practical common sense. My grateful thanks go also to Dr. Nevln Hughes-Jones, who took time to read the various drafts of the thesis, and was a constant source of encouragement. Dr. Juraj Petrik, who supervised the project, rendered exceptionally valuable assistance throughout the study, not the least remarkable of which was his deflection of an astonishing attack of the little-recognised, albeit widely-disseminated, Mad Human Disease. I acknowledge with gratitude my debt to Dr. Graeme Alexander for his helping hand in my understanding of hepatology, data presentation and, very importantly, for directing my attention to the existence of those Haydn operas.

Many people assisted my better appreciation and execution of this study. In particular, I would like to thank Dr. Voi Shim Wong, for getting me organised from the very start of the clinical phase; Gavin Pearson, for carrying out (with admirable equanimity) innumerable ELISAs and PCRs for this project; and Ayaz Majid, for a very timely suggestion. Dr. Dazhuang "John" Shang very kindly provided vital quantitation data for this study.

The following people, in various ways, enhanced the quality of my days at the EABC, thereby assisting a problem-free completion of the project and thesis: Heather Hitcham, who talked me into the project; John Saldanha, for my early training in Virology at the NIBSC; Don Hawdon and John Pegler, who, through funding by the EABC, made it financially possible to complete the project; and Mark Fletton, for cheerful and tireless assistance during the lab set-up period. Ian Reeves carried out many of the ELISAs and RIBAs, as well as providing HCV-positive blood samples. *La Clemenza di Tito* and *The Second Mrs. Kong* were but two of the animating discoveries made via the extraordinary Josephine Stubbs. The wealth of data passed on from Erik Timmers was much appreciated.

Special thanks are due Dr. Ricarda Finnern, for blazing the track, and Dr. Lorna Williamson, for her unimpeachable handling of the KOZ affair. The part played by my incredible sister Tolkins will always remain a cherished memory. Finally, I thank Jesus Christ, for being my perfect, albeit perfectly inimitable, model.

Zarah Y. Lawal
Cambridge, U.K.
May 1996

ABBREVIATIONS & CONVENTIONS

Abbreviations used in the text.

Abbreviation	Term represented
Ab	antibodies
Ag	antigen(s)
AMP	ampicillin
autorad	autoradiograph
BM	bone marrow
bp	base pairs
BSA	bovine serum albumin
BVDV	bovine viral diarrhoea virus
cDNA	complementary deoxyribonucleic acid
dATP	deoxyadenosine-5'-triphosphate (9-ß-2'-deoxy-D-ribofuranosyladenine)
DMSO	dimethyl sulphoxide
dNTPs	deoxynucleotide triphosphates (equimolar mixture of dATP, dCTP, dGTP and dTTP)
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
ELISA	enzyme-linked immunosorbent assay
EtBr	ethidium bromide
EtOH	ethanol
GBV	GB virus
GTC	guanidinium thiocyanate
h	hour
HCV	hepatitis C virus
HFV	hepatitis F virus
HGV	hepatitis G virus
HIV	human immunodeficiency virus
HLA	(human) histocompatibility leukocyte antigen
HSV	herpes simplex virus
HVR	hypervariable region
IFN	interferon
Ig	immunoglobulin
IL	interleukin
IPTG	isopropyl-ß-D-thio-galactopyranoside

IUPAC	International Union of Pure & Applied Chemistry
MAb	monoclonal antibody
min	minute(s)
mRNA	messenger RNA
nt	nucleotide
oligo	oligonucleotide
p	pico
PBL	peripheral blood lymphocytes
PBMC	peripheral blood mononuclear cell(s)
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
prog	program(me)
rATP	adenosine 5'-triphosphate (9-ß-D-ribofuranosyladenine)
RBC	red blood cell(s)
RIBA	recombinant immunoblot assay
rpm	revolutions per minute
RT	reverse transcriptase
s	second(s)
SDS	sodium dodecyl sulphate
SSC	standard saline citrate
Taq	*Thermus aquaticus*
TEMED	N,N,N',N'-tetramethylethylenediamine
temp	temperature
TET	tetracycline
TR	typing region
Tris	tirs-(hydroxymethyl)-aminomethane
U	unit
UHP	ultra-high purity
UV	ultraviolet
VZV	varicella-zoster virus
WBC	white blood cell
X-gal	5-bromo-4-chloro-3-indoyl-ß-D-galactopyranoside
YFV	yellow fever virus

IUPAC Code: Amino acids

One-letter code	Three-letter code	Full name
A	Ala	alanine
B	Asx	A or D
C	Cys	cysteine
D	Asp	aspartate
E	Glu	glutamate
F	Phe	phenylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine
V	Val	valine
W	Trp	tryptophane
Z	Glx	Q or E
-	gap	
*	end	

IUPAC Code: Nucleic acids

Symbol	Nucleic acid represented
A	Adenine
B	not A
C	Cytosine
D	not C
G	Guanine
H	not G
K	G or T
M	A or T
N	A, C, G or T

R	A or G
S	C or G
T	Thymine
U	uracil (= T)
W	A or T
Y	C or T
YFV	yellow fever virus
-	gap

INTRODUCTION

MATERIALS & METHODS

List of Figures

List of Tables

Abstract

Hepatitis C virus (HCV) is responsible for most cases of non-A, non-B hepatitis. The persistent nature of this virus has been attributed to viral replication errors, which lead to a dynamic pool of antigenic variants that allow escape from the host immune response. A major part of this escape is due to the hypervariable region 1 (HVR1) of HCV, known to encode structurally flexible, isolate-specific neutralising epitopes which undergo successive genetic alterations. In a substantial number of cases, complications of HCV infection lead to end-stage liver disease for which the only treatment is orthotopic liver transplantation (OLT). Primary HCV infection of the allograft is an almost universal phenomenon associated with OLT. This study focused on the pattern of HCV variability in the context of immunosuppression, which is a feature of post-OLT treatment. Sequences of the HCV HVR1 derived from OLT recipients and from asymptomatic (presumably immunocompetent) carriers of the virus were compared over several time-points. A rapid turnover of sequences was found in the untreated subjects, in whom mean nucleotide and amino acid sequence diversity were 19.8% and 43.5%, respectively. In the immunosuppressed patients, the corresponding figures were 2.3% and 2.3%. Untreated subjects showed a ratio of transitional to transversional mutations of 2.57, compared with 0.98 for untreated subjects ($p = 0.0165$). Similarly, the replacement to silent mutation (R/S) ratios were 8.22 and 1.33 ($p = 0.0069$), respectively. The major differences between the two groups of patients were especially demonstrated by a subset of two immunosuppressed patients, in whom the HVR1 showed almost 100% homogeneity throughout a year of follow-up. Both patients required re-transplantation within a year of the first OLT, and both died of HCV-related disease shortly afterwards. On the other hand, two other transplant recipients, who showed an HVR1 mutation rate indistinguishable from that found in the untreated group, remain well 10 months and almost two years post-OLT. This investigation (1) demonstrated that the HCV genome becomes more homogenous post-OLT; (2) suggests that, at least in some patients, the lack of genetic heterogeneity may pre-date liver transplantation; (3) highlights the critical nature of host factors in determining the clinical course of post-OLT recurrent HCV infection; and (4) provides a model of hepatitis C viral kinetics in immunosuppressed and untreated individuals. The finding of HCV genomic sequences within the leukocytes provides molecular evidence for the existence of HCV in peripheral blood cells, and thus supports the hypothesis that peripheral blood lymphocytes act as HCV reservoirs for the reinfecting virions.

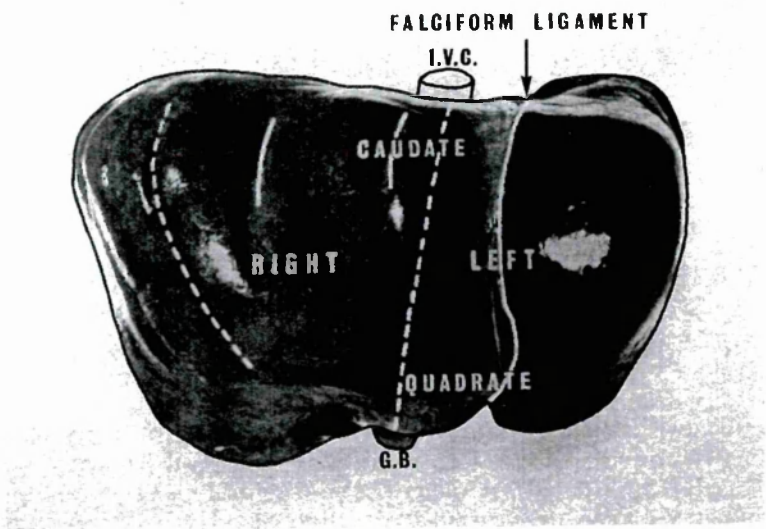
CHAPTER ONE

1.1 THE LIVER: OVERVIEW

1.1.1 Structure and function

The liver is a bilobular structure which arises from the endoderm of the developing embryo (Alberts *et al.* 1983). At 1.5 kg, or 2.5% of body weight, it is the largest solid organ of the body, as well as one of its most complex, with over 500 functions identified (Urdang and Swallow 1983). Due to its numerous metabolic functions and secretions into blood, it is classifiable as both an exocrine and endocrine gland. Like the lung, it is a segmented organ, with definite lines of cleavage between the different segments (Figure 1.1).

Fig 1.1: *Gross anatomy of the liver*



A line running through the bed of the gall bladder (G.B.) to the left margin of the inferior vena cava (I.V.C.) divides it into right and left lobes. The right lobe is sub-divided into anterior and posterior segments, and the left lobe into medial and lateral segments. Each segment is further divisible into upper and lower zones on the basis of the distributions of the hepatic artery, portal vein and bile ducts, giving 8 individual zones in theory, though boundaries are masked by the complexity of the liver.

Reproduction of FIG. 1, Chapter 1, of "Liver Disease" by A. Paton (1989), William Heinemann Medical Books Ltd., London

Liver composition includes a variety of cell types adapted to perform specific functions, such as the fibroblasts (which provide a connective tissue framework); the phagocytic Kupffer cells; the fat-storing (Ito) cells; and hepatocytes. The extracellular matrix is a complex mixture of collagens, non-collagenous glycoproteins, proteoglycans and glycosaminoglycans which, far from being inert, acts as a sensitive radar for detecting fluctuations within hepatic microenvironments (Rojkind and Greenwel 1994). The multivalent matrix macromolecules form a dynamic interface for cell-to-cell signalling. Communication between the numerous cells within the liver is essential for maintaining homeostasis, and mediators of effector functions include cytokines, growth factors and metabolites. Liver functions include carbohydrate and lipid metabolism, interconversion of nutrients, secretion of most blood plasma proteins, synthesis of bile salts, degradation of bile pigments, detoxification of chemical compounds and storage of a large number of products such as vitamins and iron (Keeton and Gould 1986). The hepatocyte, a specialised cell which derives from the embryonic gut epithelium, is the main site for these activities.

The sponge-like structure of the liver serves as a reservoir in the regulation of blood volume and flow through the body. Blood is conveyed from the gut to the liver through the portal vein, which drains the whole of the alimentary tract from the lower oesophagus to the rectum, as well as the pancreas, spleen and gall bladder. Thus the human liver, which holds at any given moment 500 ml of blood (representing a third of cardiac output and 13% of total blood supply) stands at the crossroads of the nutrient distribution system of the body. As the first organ to "see" incoming nutrients from the gut, it has a rapid enzyme induction and repression system that must adapt continually to fluctuations in nutritional intake in order to maintain proper life function.

Nutrients enter the liver in the bloodstream via the portal vein, and are distributed to the rest of the body, after processing by the hepatocytes, via the hepatic veins. Glucose is taken up and stored as insoluble glycogen, for later release when needed for maintaining normal glucose levels. Excess amino acids from protein digestion are broken down, the nitrogen removed as urea and excreted by the kidney, and the non-nitrogenous residue converted into glucose. In

addition to nutrients, liver cells receive the full force of any noxious substances that may be absorbed in the intestines.

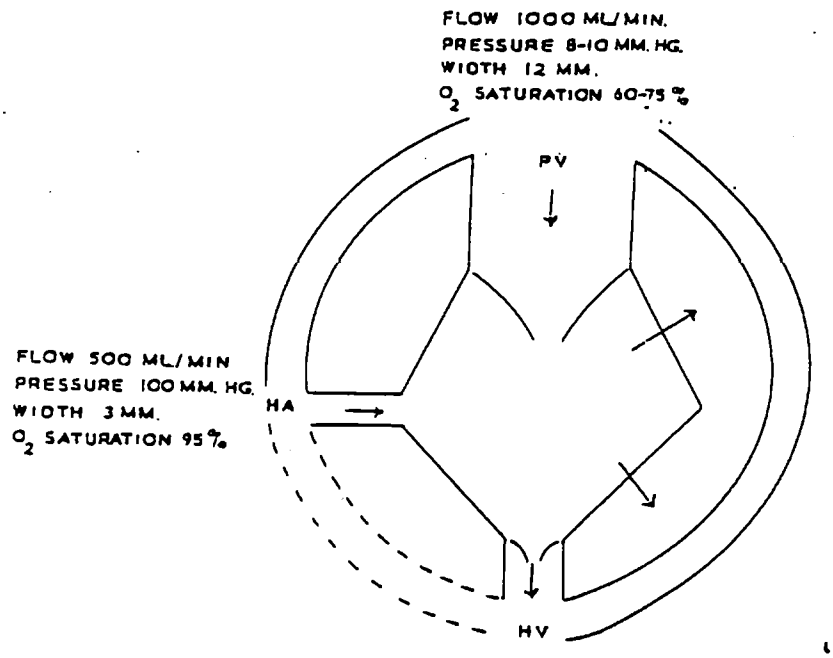
The liver is a very vascular organ, with a relatively sparse structure which is easily deformed by external forces including respiration and straining. Such forces also have an impact on blood flow. 90% of the hepatic blood is contained in the sinusoids, which form an extensive, apparently haphazard network of thin-walled spaces, the principal site for transvascular exchange between blood and hepatocytes (Paton 1969; McCuskey 1994). The functional unit of the liver, i.e., the minimum mass of tissue required to perform hepatic function, is the hepatic lobule. This is defined as a cone-shaped array of cells and interconnecting sinusoids bounded at its convex surface by terminal branches of portal venules and hepatic arterioles, and with its apex at the central vein (McCuskey 1994). Fully-oxygenated arterial blood reaches the sinusoids via the high-pressure (100 mmHg) hepatic arterial system as well as the low-pressure (8-10 mmHg) portal vein (Figure 1.2). The pressure from arterial blood flow is presumably dissipated during its spread through the sinusoids. The complex nature of intrahepatic blood is dependent on multiple factors, including genetic influences on the sinusoidal arrangements as well as specific circulatory adjustments. Microvascular injury leads to an inflammatory response manifested by leukocyte adhesion to the endothelial lining of venules and sinusoids. Platelet adherence is also found in more severe injury, causing plugging of the blood vessel as well as a decrease or cessation of blood flow.

Among the most fascinating features of the liver is its impressive capacity for regeneration, with restoration to normal size even after 80% removal reported in humans and animals (Paton 1969). Furthermore, this capacity does not diminish with age. Control of the various kinds of liver cells and extracellular matrix is as important as maintaining the dynamic equilibirum between the bloodborne products of liver metabolism.

The ability of the liver to withstand insults is determined by many interlinked factors, both external and internal, of which some are given in Table 1.1. Hepatic damage usually induces a short-lived, reversible illness. Despite this limited response, a wide range of functional

derangements occur. It is difficult to find a direct "cause and effect" toxic action of the liver, divergent symptoms of hepatitis probably arising from the variable interplay between the different factors. In a few individuals, liver damage leads to progressive and irreversible liver disease.

Figure 1.2: *Scheme of liver sinusoid showing relative contributions of hepatic artery (HA) and portal vein (PV). HV = hepatic vein.*



Reproduction of FIG, 2, Chapter 1, of "Liver Disease" by A. Paton (1989), William Heinemann Medical Books Ltd., London

A central question in chronic liver disease is whether the initial damaging agent is responsible for perpetuating the injury, or whether it merely sets off a self-damaging (immunological) reaction on the liver. At one end of the spectrum, alcoholic cirrhosis results from repeated assaults in susceptible individuals. In primary biliary cirrhosis, at the other end, immunity against glycoproteins from injured bile ducts is the major identifiable cause of liver damage, but may merely reflect an aggravated reaction to the original acute injury. The reason for continued damage is unknown. Inflammation may prevent proper nutrition of hepatocytes at the sinusoidal level. Another possibility is the activation of Kupffer cells to produce antibodies against the original antigen.

Table 1.1: Factors influencing liver function

External	Internal	Genetic factors
Nutrition	Blood flow	Species
Infectious agents	Blood pressure	Enzymes
Alcohol	Oxygen supply	Anatomical peculiarities
Poisons	CO_2 concentration	Age
Drugs	Temperature	Sex
Antigen/antibody	Nutrient supply	
	Pregnancy	

1.1.2 Causes of hepatic disease

Hepatitis, an inflammatory condition of the liver first described by Hippocrates, may be caused by bacteria, parasites, drugs, alcohol, toxins and transfusions of incompatible blood, as well as by viruses (Urdang and Swallow 1983). Table 1.2 shows some common and rare etiologies of liver disease. Symptoms may be mild and brief, or fulminant and life-threatening. A litre of bile is produced daily by the human liver, and comprises mainly conjugated bilirubin and bile salts, phospholipids and cholesterol. Bilirubin is produced at the rate of about 300 mg/day, mostly as a breakdown product of haemoglobin in the reticuloendothelial system. Most is excreted in the stools as stercobilin. Obstruction of bilirubin excretion, or the occurrence in the liver of bilirubin loads greater than it can deal with, may lead to jaundice. In this diseased condition, the eyeballs, skin and urine become abnormally yellow as a result of bile in the blood.

Transaminases such as alanine aminotransferase (ALT) and aspartate aminotransferase (AST) are enzymes responsible for the transamination of amino acids. Found throughout the body, they are particularly abundant in muscle, especially heart muscle, and liver. Serum transaminase (most commonly ALT) levels are used as a screening test for liver dysfunction. Normal ALT levels in blood range from 7-50 international units (IU)/l, and may be as high as 500 - 5000 IU/l in acute viral hepatitis. However, a rise in serum ALT does not

Table 1.2: *Etiologies of liver disease*

Etiology	Selected agents or specific defects
Infections	
- Acute viral infections:	HAV; HBV; HCV; HDV; HEV; CMV; EBV; HSV-1; HIV; HFV; HGV; VZV; mumps virus; yellow fever virus
- Chronic viral infections	HBV; HCV; HDV; HGV
- Fungi	Candida; cryptococcus
- Parasites	Amoeba; schistosomia
- Bacteria	Salmonella; brucella; legionnella
- Mycobacteria	*M. tuberculosis*; *M. leprae*
- Spirochetal	*Leptospira interrogans*; *Trepanema pallidum* (syphilis agent)
Drugs and toxins	
- Ethanol	
- Chemicals	Be; CCl_4; As; P
- Heavy metals	Fe; Cu
Physical agents	Heat; hypoxia; ionising radiation
Genetic disorders	Cystic diseases; Wilson's disease
Metabolic disorders	Porphyrias; metal storage disorders
Nutritional diseases	Kwashiorkor
Pregnancy	
Immunologic/unknown etiology	Primary biliary cirrhosis; autoimmune hepatitis

Adapted from Table 5.1 in "Diseases of the Liver and Biliary Tract: Standardization of Nomenclature, Diagnostic Criteria and Prognosis" - International Hepatology Informatics Group, Leevy et al. (eds), Raven Press, New York 1994

necessarily reflect a purely hepatic origin nor indeed signify liver damage. Hence neither serum ALT concentration nor the occurrence of jaundice can reliably be taken by itself as an

indicator of liver function. Other non-specific symptoms of liver dysfunction include abdominal discomfort, vomiting and anorexia. More unequivocal signs include necrosis, chronic liver damage and hepatoma.

Chronic hepatitis is defined as liver disease of viral or any other aetiology that persists for more than six months. Increasingly severe hepatitic states are classically designated chronic persistent hepatitis (CPH) and chronic active hepatitis (CAH). These are characterised by varying degrees of liver cell damage, starting with swelling and granularity of hepatocytes. As disease advances, blood flow through the sinusoids is progressively impaired by the swelling of endothelial and Kupffer cells. The sinusoidal lumen is further constricted by attachment of granulocytes to the endothelium. Other impedances to blood flow include adherence of lymphocytes and platelets to Kupffer cells, and the inflow of extrahepatic monocytes into the sinusoids. Eventually, lesions are produced in which sinusoids are obliterated by swollen parenchymal cells (hepatocytes). With continuous activity (persistent hepatitis), fibrous tissue may extend into the lobule. Patchy (piecemeal) necrosis may occur at the periphery of the lobule, leading to groups of liver cells surrounded and isolated by inflammatory infiltrates. Some hepatic lesions, such as hepatocyte enlargement with sinusoidal encroachment, are reversible. Others, such as fibrosis, are irreversible.

The term chronic active hepatitis defines persisting activity of the hepatic inflammatory process, which is "aggressive", as opposed to the subdued course of continuing damage which occasionally follows virus infection. Chronic liver disease of viral or any other aetiology can lead progressively to the development of liver fibrosis and ultimately cirrhosis. Cirrhosis is defined as a chronic illness characterised by fibrosis that distorts the liver architecture, in the presence of generalised damage and the formation of regeneration nodules (McCuskey 1994). Ito cells have been identified as the main source of the fibrillar matrix proteins observed in liver fibrosis (Arthur 1994). While the precise mechanism of liver fibrogenesis remains to be identified, the most likely routes are via T-cell mediated Kuppfer cell activation or immune-mediated hepatocyte damage. Hepatic function is crucially dependent on the maintenance of a proper balance of cells and extracellular matrix by the various metabolites, cytokines and

growth factors. Disruption of this delicate ecosystem causes functional derangement. Damaged hepatocytes are known to release a factor which is mitogenic specifically for the hepatic lipocytes, transforming them to the fibrogenic myofibroblasts.

Methods for determining the extent of liver damage include radiology, hepatic ultrasonic scanning and histologic analysis of needle biopsies. As advances in the understanding of liver histopathology have been made, the need has arisen for a diagnostic system which more accurately portrays the complex stages of chronic hepatitis. Classification schemes reflecting the etiology, the extent of inflammation (grading) and the degree of progression (staging) have been proposed. The most widely used of these is the histological activity index, also known as the Knodell score, which consists of four separate scores for different components of the lesions (BegicJaneva and Boricic 1995; Callea *et al.* 1995). An expanded system incorporating additional components of liver pathology in chronic HCV infection was recently proposed (Wong *et al.* 1996).

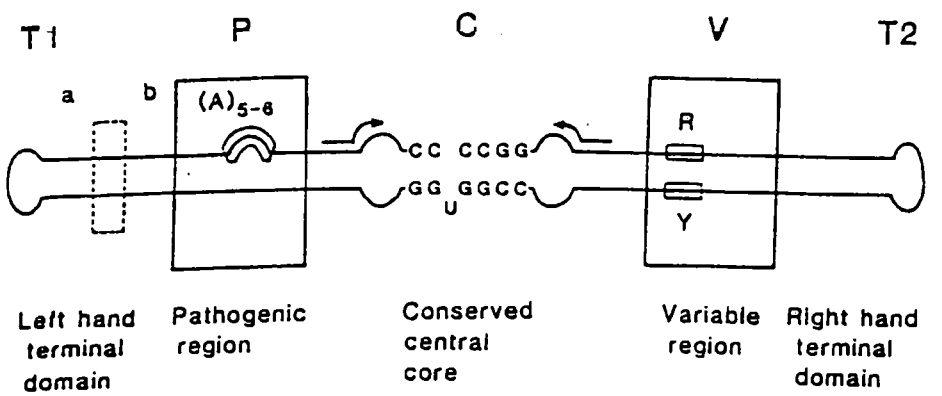
1.2 THE ROLE OF VIRUSES IN LIVER PATHOGENESIS

1.2.1 Definitions

Viruses are infectious agents which have a nucleic acid (DNA or RNA) genome (Jawetz *et al.* 1974) enclosed by a symmetric protein shell, termed a **capsid**. The unit thus formed constitutes the **nucleocapsid (or core)**. The **virion**, which comprises the viral nucleic acid enclosed in a protein coat, composes a single infectious unit, and is identical with the nucleocapsid in some instances (e.g. picornaviruses). In more complex viruses, the virion includes the nucleocapsid plus a surrounding envelope. The **primary nucleic acid structure** refers to the sequence of bases in the nucleic acid chain, and the **secondary structure** to the spatial arrangement of the complete nucleic acid chain (e.g., whether circular or linear). Other elements of fine spatial detail in the helix (such as presence of supercoiling, breakage points, catenation, gaps, regions of strand separation) contribute to its **tertiary structure.**

A class of subviral pathogens, termed **viroids**, was discovered in the 1970s. The smallest known class of autonomously replicating pathogens, viroids are single-stranded, circular RNA molecules of 246-375 residues with a unique secondary structure characterized by double helices separated by short, unpaired stretches (Riesner *et al.* 1979). Unlike virions, where the infectious unit includes the surrounding coat, the viroid RNA by itself constitutes the infectious agent. Common viroid features include the number of base pairs, high co-operativity and formation of very stable hairpins. These characteristics, found to be improbable in random sequences, imply that the various viroid species, although differing in nucleotide sequence, are structurally very similar. There is no evidence of any mRNA activity in viroids, implying that all required enzymic functions reside in the plant hosts. Despite this apparent lack of protein coding capacity and the small size of viroids, they appear to be multigenic, with functions corresponding to structurally distinct, interchangeable domains (fig 1.3).

Figure 1.3: *Model of five viroid domains (T1, P, C, V and T2):*



Arrows depict an inverted repeat; a and b are segments of the T1 domain that flank a region of RNA exchange (dashed box); R and Y indicate a short oligopurine-oligopyrimidine helix.

Reproduction of Fig. 1, Keese & Symons (1985).

The unique structural features of viroids may represent an optimal compromise between stability and flexibility. The native viroid structure is completely stable below its sharp transition temperature of 49°C but opening of the structure, a necessary prerequisite for replication, is easier to induce in viroids than in comletely double-stranded RNA. On the other hand, the extensive base-pairing to give rod-like structures would allow a certain degree of protection against degradation.

Ribozymes are RNA molecules capable of cleaving RNA target molecules in a catalytic fashion (Lyngstadaas *et al.* 1995). The hammerhead ribozyme (HHR), one of the smallest RNA enzymes known, has been extensively investigated due to its potential as an antiviral agent. It is a small RNA motif comprising three base-paired stems connected by conserved sequences which are essential for catalysis. Distinguished by the "Y" shape of the three stems, HHRs are stabilised in the central core by a network of hydrogen bonds involving, in particular, two non-Watson-Crick G:A pairs (Laugaa *et al.* 1995). According to Ohkawa *et al.* (1995), HHRs belong to a class of molecules known as antisense RNAs. Due to short extra sequences that form the so-called catalytic loop, they can act as enzymes. Since the catalytic domain captures Mg2+, which can cleave phosphodiester bonds, HHRs are recognised as metalloenzymes. Autocatalytic ribozymes have been described in both natural and synthetic RNAs (Williams *et al.* 1995; Ekland and Bartel, 1995). The genome of human hepatitis D virus (HDV) has been shown to have ribozyme properties.

An open reading frame (**ORF**) is a contiguous stretch of 50 or more codons uninterrupted by stop signals. The polarity of viral nucleic acid strands is defined such that the mRNA has **plus (+) polarity**. The protein-encoding strand of HBV DNA, which is transcribed into mRNA, therefore has **(-) polarity**, while the same strand in the RNA HCV genome has (+) polarity. **Retroviruses** can convert information in their RNA genomes to DNA using the virally-encoded enzyme **reverse transcriptase**. Reverse transcription of viral RNA into DNA is an essential step of the genomic replication in hepadnaviruses such as HBV, which have DNA genomes

and are termed **pararetroviruses**, in contrast to **orthoretroviruses** such as HIV, which have RNA genomes.

1.2.2 Virus Taxonomy

The earliest efforts to classify viruses relied on filterability as the single physicochemical characteristic measurable, and classification was based on such features as common pathogenic properties. Viruses which shared the pathogenic property of causing hepatitis (HAV, HBV, YFV, and Rift Valley fever virus) were thus grouped together as the "hepatitis viruses."

The Baltimore classification scheme (Baltimore 1971) focuses on the central role of mRNA in virus multiplication, organising the myriad animal viruses according to their modes of transcription and replication. According to this scheme, six classes of animal viruses are defined by the structural relationship between their mRNA and the nucleic acid in the virion:-

Class I: double-strand DNA viruses (e.g. vaccinia), where the double-stranded DNA genome acts as template for RNA synthesis, as is the case with cells.

Class II: single-strand DNA viruses. If the mRNA is the same strand as the DNA, the intermediate sunthesis of a (-) strand DNA is needed to provide a template for mRNA synthesis.

Class III: double-strand RNA, e.g. reovirus.

Class IV: positive-strand RNA viruses (e.g. HCV), where the genome RNA and mRNA have the same polarity.

Class V: negative-strand RNA viruses - the virion RNA has a base sequence complementary to that of the mRNA.

Class VI: retroviruses, which direct the formation of DNA that acts as the template for making their RNA.

In the formal taxonomic system (International Committee on Taxonomy of Viruses or ICTV) applied universally today, viruses are classified into families on the basis of several parameters, the type and form of the nucleic acid genome (DNA or RNA, single- or double-

stranded) being the most important (Jawetz *et al.* 1974). Next in significance are the size and morphology of the viral particle, including the type of symmetry (icosahedral or helical) and the presence of membranes.

Virus families, designated by terms with the suffix "*-viridae,*" represent populations of viruses that share common characteristics and are distinct in many ways (Murphy and Kingsbury 1991). Subfamilies, whose designations are given the suffix -"*virinae,*" have been introduced for four families (*Poxviridae*, *Herpesviridae*, *Papovaviridae* and *Retroviridae*) to reflect the intrinsic complexity of their intrafamilial relationships. Within each family, virus genera represent subsets of species that share common characteristics. Terms representing genera are suffixed with "*-virus.*" The term "species" refers to individual viruses or clusters of viral strains with many features in common, and is currently defined to include the vernacular term "virus."

1.2.3 The compass of viral hepatitis

Viral hepatitis is a systemic disease primarily involving the liver, which can be caused by several agents. Clinical manifestations may also vary from one host to another. The most common types of acute viral hepatitis in humans are hepatitis A (infectious or short-incubation hepatitis), caused by hepatitis A virus (HAV); hepatitis B (long-incubation or serum hepatitis) which results from hepatitis B virus (HBV) infection; and hepatitis C, associated with the recently discovered hepatitis C virus (HCV). HDV is a subviral particle which requires coinfection with HBV for its own expression and replication, and HEV, like HAV, is usually transmitted enterically. Less well-characterised hepatitis viruses are hepatitis F virus (HFV), believed to cause acute sporadic hepatitis; and the chronic hepatitis agents HGV and the GB viruses. While HAV and HEV cause only acute hepatitis, HBV, HCV and HDV frequently establish persistent infections, which with time can cause chronic inflammation and liver cirrhosis. Variants of the hepatitis viruses have been identified in infected hosts. Although naturally occurring mutations have been found in all viral genes, they are usually concentrated in genes encoding the structural envelope and nucleocapsid

proteins. The resulting viral variability is important as different mutations may correlate with different disease profiles. For example, specific mutations may impede viral clearance by providing escape variants from host immunity or antiviral therapy (Blum 1995; Coppola and Rizzetto 1995).

Hepatitis is an occasional feature of the clinical syndromes characteristic of several other viral infections (White and Fenner 1986). Thus, yellow fever virus (YFV), a flavivirus, causes the very severe jaundice from which it derives its name. Hepatitis is also prominent in many of the haemorrhagic fevers, such as Lassa, Ebola and Cremean-Congo haemorrhagic fevers. It is a major feature of the widespread viral infections which overwhelm neonates or immunocompromised patients - (neonatal) herpes simplex and cytomegalic inclusion disease caused, respectively, by HSV and CMV (Brooks *et al.* 1991; Kosai *et al.* 1991). HIV infection is frequently associated with viral hepatitis, which may alter the course of infection (Duffy *et al.* 1986; Schneiderman *et al.* 1987; Kahn *et al.* 1991). Other well-characterized viruses that infrequently cause sporadic hepatitis are Epstein-Barr virus (EBV), rubella virus and the enteroviruses (Okano *et al.* 1991; Sugaya *et al.* 1988; Read *et al.* 1985). Occasionally, systemic diseases such as tuberculosis and gonorrhoea may have their principal clinical manifestation as liver injury. The pregnant state appears to make women more susceptible to hepatitis, perhaps due to aggravation of portal hypertension, or subtle changes in handling of bilirubin in the third trimester of pregnancy. Although pregnancy is not significantly associated with a worse outcome of most viral infections, a clear association between HEV-related fatalities and pregnancy has been shown, and this may be the case with other, identified and as yet unidentified, viruses. Liver injury associated with alcoholism may be related to acetaldehyde toxicity and/or nutritional deficiency, and may also be triggered by viral infection in genetically susceptible individuals. Finally, as yet undiscovered viruses may be at the heart of many cryptogenic liver diseases such as primary biliary cirrhosis (PBC) and autoimmune chronic hepatitis (AIH), which are currently considered to have an autoimmune basis.

1.3 MAJOR AGENTS OF VIRAL HEPATITIS

Viral hepatitis may be as old as recorded history (Alter and Seeff 1993). References from the Middle ages are plentiful, with some dating back to the biblical period. The unsanitary conditions surrounding recorded outbreaks of epidemic jaundice suggest that these descriptions referred to infection with HAV, or even HEV, both of which are oral-faecally transmitted. The origin and evolution of of the transfusion-related HBV and HCV are unknown, although these may also have existed for centuries, perhaps transmitted by insect vectors. The incidence of transmission of both viruses has probably increased in the wake of modern technological advances such as use of the hypodermic needle, and blood transfusion. Today, viral hepatitis occurs in sporadic outbreaks as well as epidemics. Most viruses produce characteristic histopathological lesions in the liver. Although these are generally indistinguishable, specific serological tests enable the identification of individual agents. Some characteristic features of the most common hepatitis viruses, described in the following pages, are summarised in Table 1.3.

Table 1.3: *Features of the most common viral hepatitis agents*

Virus	Nucleic acid type	Family	Env-elope?	Size (nm)	Genomic length (kb)	Chronicity
HAV	RNA	*Picornaviridae*	No	27	7.5	None
HBV	DNA	*Hepadnaviridae*	Yes	42	3.2	5%
HCV	RNA	*Flaviviridae*	Yes	30-60	10.0	>80%
HDV	RNA	Unclassified	Yes (HBV-derived)	42	1.7	Coinfection <10%; Superinfection: >80%*
HEV	RNA	*Caliciviridae*	No	33	7.8	None
HGV	RNA	*Flaviviridae?*				

* see section 1.3.3

1.3.1 Hepatitis A virus (HAV)

This is a single-stranded, nonenveloped RNA virus of positive polarity, classified as a member of the *Picornaviridae* family, which was originally placed in the genus *Enterovirus* as enterovirus 72 (Robertson *et al.* 1991). The virion consists of the linear RNA genome and a protein shell comprising 4 proteins, VP1 - VP4 (Weitz and Siegl 1993). Although extensive homologies in protein structure exist with poliovirus, there is little nucleotide sequence homology between HAV and other members of its family, which led to its re-classification into a separate genus, 'hepatovirus' (Minor, 1991). Its 7500-nucleotide genome contains a single open reading frame (ORF) which runs between nt 734 and nt 7415, and is flanked by untranslated regions at the 5' and 3' ends. The coding region can be subdivided into three regions: P1, which specifies the structural proteins VP1-4; P2, which codes for three nonstructural proteins (2A-C); and P3, which encodes four additional nonstructural proteins (3A-D).

One of the most distinguishing features between classical picornaviruses and HAV concernns replication: While viruses such as poliovirus or human rhinovirus replicated to high titres in susceptible cells, with replication cycles of 8 - 18 h, HAV replicates slowly and inefficiently in cell culture (Siegl *et al.* 1984). Adaptation to culture may take months, and viral titres usually remain low therafter. It is also remarkable that HAV, unlike poliovirus, fails to shut down the host cell metabolism, leading to persistent infection in culture (de Chastonay and Siegl 1987).

Human HAV strains have been classified into four genotypes (I, II, III and VII) which, despite a 15-25% difference in nucleotide sequence of the P1 region, define only one antigenic serotype, so that a single infection can confer lifelong immunity. Three additional genotypes (IV, V and VI) are unique simian strains which have significant antigenic differences from the human strains (Lemon *et al.* 1992).

HAV has a worldwide distribution, although its prevalence varies in different geographic locations, probably reflecting socioeconomic conditions. In developing countries such as Ethiopia and Brazil, over 90% of children are infected by 10 years of age. This contrasts with a figure of less than 5% prevalence by 25 years of age in Sweden (Zachoval and Deinhardt 1993).

HAV has a distinct resistance to antiviral agents, such as 2,4-dichloropyridine and guanidine, which are effective with other picornaviruses (Weitz and Siegl 1993). The strength of interaction between the capsid units and the viral RNA is the probable cause of this effect, and is also reflected in the high thermal stability of HAV: in the presence of 1 M magnesium chloride, it retains its structural integrity and biological functions even at temperatures up to 80°C. Infection may, however, be prevented by sanitation, and by a timely administration of normal human immunoglobulin (NHIG) to individuals exposed to the virus. Killed hepatitis A virus vaccines are based on the demonstration that formalin-inactivated virus extracted from the liver of infected marmosets induced protective antibodies on challenge with live vaccine (Provost and Hilleman 1975). Current live attenuated vaccines require administration by injection. Advantages of such vaccines include ease of administration, relative low cost and long-term protection. Acute hepatitis A usually follows a benign course, and there is no documented human chronic carrier state or animal reservoir for HAV. The potential for eventual eradication of this virus from human populations is therefore high.

1.3.2 Hepatitis B virus (HBV)

HBV, the prototype for the family *Hepadnaviridae*, has an unusual genome - a small, partly double-stranded DNA molecule with a single-stranded region of variable length (Tiollais *et al.* 1985). The long, or L(-) strand is linear and of fixed length, while the short, or S(+) strand, also linear, varies in length from isolate to isolate. At 3200 bp, it is the smallest of any known human DNA virus (Carman *et al.* 1993a). HBV infection is clinically important because of its high frequency of occurrence, often complicated by HDV superinfection (Sonigo and Tiollais 1985). The demonstrated presence of free and integrated forms of HBV DNA in

infected host leukocytes (JianRen, *et al.* 1985) indicates the possible role of these cells in promoting the viral persistence which forms a frequent feature of HBV infection. Integration of the HBV genome within hepatocytes has been postulated as a trigger for the hepatocellular carcinoma with which this virus is significantly associated. The production of new vaccines by recombinant DNA techniques has been made possible by advances in understanding mechanisms of viral replication and integration in hepatocellular carcinoma (Tiollais *et al.* 1985). This and other aspects of HBV biology are discussed more comprehensively in Section 1.4.

1.3.3 Hepatitis delta virus (HDV)

The delta agent (HDV) is an unusual human virus first described in 1983 (Rizzetto 1983) which has many similarities to pathogenic RNAs of plants. Its many intriguing features include a circular, 1.7 kb RNA genome, a characteristic unique among animal viruses (MacNaughton *et al.* 1993); and the requirement of the envelope from a helper hepadnavirus for hepatitis D viral particle assembly. Cultured cells co-transfected with a cDNA clone of HDV RNA genome and cDNA for woodchuck hepatitis virus (WHV, a hepadnavirus) were reported to release infectious D particles (Netter *et al.* 1995). In humans, the helper virus is usually HBV (De Wit and Coutinho 1983; Samuel *et al.* 1995).

The single-stranded RNA genome has a high (60%) G+C content, and has the ability to fold on itself by intramolecular base pairing to form an unbranched rod structure (Mason and Taylor 1989; Saldanha *et al.* 1990). Under the electron microscope, HDV RNA molecules in the native conformation appear as compact rods, which convert to circular molecules under denaturing conditions (Kos *et al.* 1986). Due to its double-strandedness under native conditions, HDV RNA is relatively stable under a variety of manipulations. There is evidence that it can act as a ribozyme, autocaytically cleaving and ligating itself by an unusual mechanism (Bartolome *et al.* 1995). The circularity, high degree of intramolecular base pairing and mode of replication of HDV suggest a close resemblance to viroids, discussed in section 1.2.1.

The viral particle is composed of the HBV-derived envelope (HBsAg) and a nucleocapsid comprising hepatitis delta antigen (HDAg) and the RNA genome (Lai 1995), both of which may play crucial roles in the viral life cycle. There are three structural regions, which comprise: the ORF for HDAg, which is the only encoded protein; a region with ribozyme activity; and a heterogenous region with no known structure or function. HDAg specifically binds both the genomic and antigenomic strands of the HDV RNA (Poisson *et al.* 1993), and is the basis of serological assays for the virus (Negro and Rizzetto 1995). The 0.8 kb antigenomic strand is postulated to be the mRNA for delta antigen (Polo *et al.* 1995), of which two forms appear sequentially in infected cells - the small form, which positively regulates HDV replication, and the large form, a negative regulator of replication which is essential for particle assembly and virion export (Cullen *et al.* 1995). At least three types of HDV have been identified, which have different pathogenicities and geographic distributions (Casey *et al.* 1993). Most published sequences, including those from North America, Europe, the Middle East and the South Pacific belong to a single genotype. A second genotype is represented by a single Japanese isolate. Sequences from S. America constitute a third genotype, associated with a particularly severe form of type D hepatitis

The delta agent is frequently found among drug addicts, haemophiliacs and multiply transfused patients, risk factors shared with HBV (Jilg *et al.* 1984). The importance of HDV resides in its ability to convert an asymptomatic or mild, acute or chronic HBV infection into a fulminant or severe, progressive disease, characterized by massive liver necrosis (Vigevani *et al.* 1985.) Simultaneous transmission of HDV with HBV (coinfection) usually leads to a clinical picture similar to that seen with hepatitis B alone. However, superinfection of a chronic HBsAg carrier results in a severe, sometimes fatal, acute hepatitis, followed by chronic liver disease in up to 40% of cases (Genesca *et al.* 1987; Jilg *et al.* 1985). HDV superinfection has also been linked with hepatocellular carcinoma (Cronberg *et al.* 1984; Wu *et al.* 1995). In addition to intravenous drug abuse and household contact with HBsAg-positive carriers, heterosexual activity seems to be an efficient route of HDV transmission

(Stroffolini *et al.* 1994). In the absence of a specific therapy for delta agent infection, vaccination against hepatitis B is used as a prophylactic measure (Jilg *et al.* 1984).

1.3.4 Hepatitis E virus (HEV)

The first documented evidence of ET-NANBH came from New Delhi in 1955 (Purcell and Ticehurst 1988). It is a sporadic, epidemic form of acute viral hepatitis also found in Africa (Belabes *et al.* 1988; Bradley *et al.* 1993) and in the Indian subcontinent, South East and Central Asia, and N. America (Bradley *et al.* 1988; Bradley *et al.* 1993). Unlike HAV, it affects predominantly young to middle-aged patients. Always self-limiting, its clinical profile is similar to that of hepatitis A, with the striking distinction of a 20% case mortality rate in pregnant women. This unexpectedly high fatality rate may reflect a unique feature of viral replication or liver pathogenesis in these individuals. Identification by immune electron microscopy of the etiological agent (Humphrey *et al.* 1988) was followed by its further characterization and designation in 1990 as hepatitis E virus (HEV) (Reyes *et al.* 1990). HEV is responsible for more than 50% of symptomatic acute hepatitis cases in developing countries (Purdy and Krawczynski 1994). This is consistent with the postulate of an enterically transmitted agent because environmental conditions in the most endemic areas favour a faecal-oral transmission route.

HEV is a spherical RNA virus 27-34 nm in diameter, with a morphology and biophysical properties most similar to those of the family *Caliciviridae* (Wattre 1994; Bradley 1995). Its genome comprises an approximately 7.5 kb positive-sense, single-stranded RNA molecule with three distinct ORFs (ORF1, ORF2 and ORF3). Isolates can be broadly divided into two serotypes, represented by isolates from Mexico (HEV-M) and Burma (HEV-B), of which HEV-B appears to be more common. Genetic analysis of cDNA clones representing nearly the entire genome revealed nucleotide sequence variation, including an HVR region in ORF1. A nucleotide sequence diversity of 58% has been reported between the two serotypes in this region. Cross-challenge studies show, however, that animals infected with HEV-B are immune from reinfection by HEV-M.

An additional peculiarity of this virus is the relatively low frequency of clinical disease which develops in case contacts, an unusual feature for an enterically-transmitted agent: A 2.4% HEV transmission rate was reported during an epidemic in Nepal, which contrasted with the 10-20% rate observed in household contacts of patients with hepatitis A occurring in the same region (Bradley *et al.* 1993).

1.3.5 Hepatitis F virus (HFV)

A novel agent, designated HFV, was recently reported to be the cause of a sporadic form of non-A, non-B hepatitis (Deka *et al.* 1994). Described as comprising 27 - 37 nm particles, the presumed agent was serially transmitted in rhesus monkeys by intravenous inoculation of the stool extract from a French patient. The genetic material of HFV is a 20-kb, double-stranded DNA molecule which, upon partial sequence analysis, was found to resemble none of the other enterically transmitted viruses. A more definitive description awaits the sequencing of the entire genome.

1.3.6 Hepatitis G virus (HGV) & the GB Viruses

10 - 20% of non-A, non-B hepatitis is caused by non-C, non-E etiologic agents. One potential agent, associated with serum from a surgeon (GB) who developed acute hepatitis and designated the GB agent, was able to induce hepatitis in experimentally infected tamarins. Two novel flavi-like viruses, designated GBV-A and -B, have recently been cloned and characterized from the GB agent (Schlauder *et al.* 1995; Simons *et al.* 1995). A third agent, designated GBV-C, was isolated from a different human specimen (Simons *et al.* 1995). GBV-A, -B and C share limited overall sequence identity, but have a genomic organisation similar to other flaviviruses. An additional RNA virus, designated hepatitis G virus (HGV), was recently isolated from the plasma of a patient with chronic hepatitis (Linnen *et al.* 1996). HGV, which has a 9392-nucleotide genome encoding a putative polypeptide of 2873 amino acids, is closely related to GBV-C, and distantly related to GBV-A and -B. A transfusion-transmissible agent, it has a global distribution, and is closely associated with both acute and chronic hepatitis.

1.3.7 Hepatitis C virus (HCV)

The single-stranded RNA genome of HCV contains a large translational open reading frame (ORF), which encodes a polyprotein of about 3,000 amino acids. The main cause of parenterally-transmitted hepatitis, it is an enveloped virus 30-60 nm in diameter (Plagemann 1991). Phylogenetic analysis shows a strong resemblance between HCV and both flaviviruses and pestiviruses, leading to its current classification as a separate genus within the family *Flaviviridae* (Weiner *et al.* 1991). Although hepatitis C is is often mild, HCV infection is remarkable for its high frequency of progression to chronicity (over 50%), with end-stage liver disease developing in up to 20% of carriers (Alter 1988). HCV is described in detail in Section 1.5.

1.4 HEPATITIS B VIRUS: A MODEL HEPADNAVIRUS

1.4.1 Historical background

Two types of human hepatitis were recognised before the 1960s, distinguishable by their modes of transmission - enteric (type A) and parenteral (type B) - and designated hepatitis A and B, respectively. In 1963, a search for polymorphic serum proteins led to the discovery, in the blood of an Australian aborigine, of an unknown antigen - later named the Australia antigen (Blumberg *et al.* 1967). Immune electron microscopy studies by D.S. Dane showed that the antigen was present on the surface of particles with three different morphological forms and was associated with the disease serum hepatitis, now known as hepatitis B (Dane *et al.* 1970). The 22-nm Australia antigen particles, subsequently renamed hepatitis B surface antigen (HBsAg), were found to be noninfectious, but the 42-nm particles were shown to be infectious virions containing an endogenous, DNA-dependent DNA polymerase within their cores (Kaplan *et al.* 1973). The virus, named hepatitis B virus (HBV), was subsequently characterised in 1974 as a small, circular DNA which was partially double-stranded and partially single-stranded (Robinson and Greenman 1974a; Robinson *et al.* 1974b). Other HBV-like agents have been discovered in the woodchuck, ground squirrel, Pekin duck

and heron. Together with HBV, these viruses comprise the virus family *Hepadnaviridae*, a name derived from their hepatotropism and DNA genome (Gerlich 1993).

1.4.2 Structure and molecular biology

The HBV virion is a spherical particle 42 nm in diameter, consisting of a 27-nm icosahedral core within a closely adherent "envelope" containing virus-specific surface antigen (HBsAg). The core contains two additional antigens, HBcAg and HBeAg, plus DNA polymerase and viral nucleic acid. The genome consists of a circular, double-stranded DNA with a molecular weight of about 2.2 x 10^6, a length of 3200 nucleotides and a most unusual structure: One of the DNA strands is incomplete, leaving 15 - 50% of the molecule single-stranded, but the gap is repaired *in situ* by a DNA polymerase carried in the core (White and Fenner, 1986; Collier and Oxford, 1993).

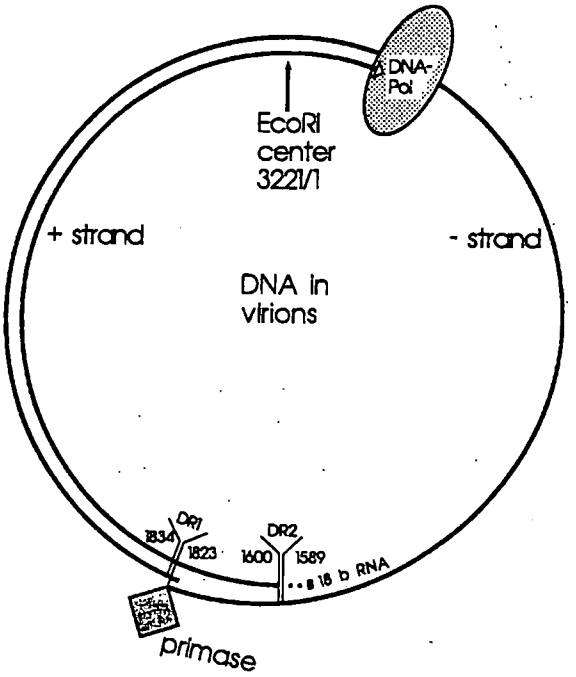


Figure 1.4: *Structure of HBV DNA*

Numbering begins at a cleavage site of EcoRI. *DR1 and DR2 represent two directly repeated sequences.*

Reproduction of Fig. 6.3, p.90, Chapter 6, Section 2, of "Viral Hepatitis: Scientific basis and clinical management," by A. J. Zuckerman & H.C. Thomas, eds. (1993), Churchill Livingstone, Edinburgh

The complete (long, or (-)) strand contains a discontinuity at a unique site, and has attached to its 5' end a covalently bound protein (primase), which primes (-) strand synthesis (fig. 1.4). It also has a terminal redundancy of 9 bases, where the ends ovelap slightly. The (+) strand is connected with the HBV DNA polymerase at its 3' end. Its 5' end is formed by 18

ribonucleotides, which are capped in the same way as mRNA. There is also a short triple-stranded region at the point where the (+) strand bridges the discontinuity of the partly overlapping (-) strand ends.

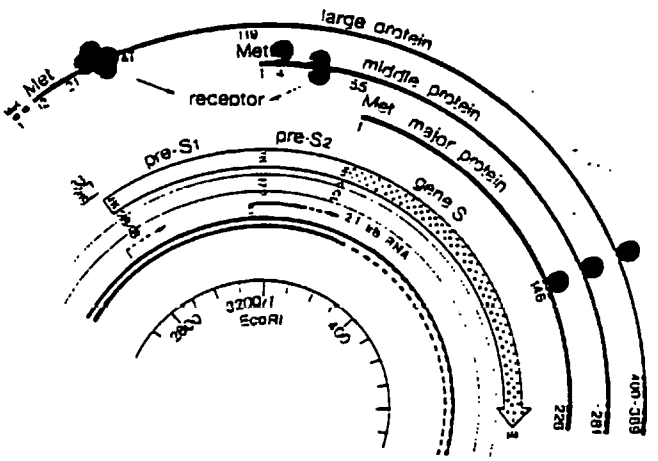
HBV shows two deviations from the usual classification criteria: it contains both DNA and RNA, and its genome contains partly single-stranded, partly double-stranded and even triple-stranded DNA. The mechanism of HBV attachment to and penetration of host cells is unknown. Even less is known about the way in which the viral genome is released within the cell.

A fundamental difference between hepadnaviruses and retroviruses lies in the fact that the form of the genome present in mature virus particles is DNA, and not RNA. Hepadnaviruses replicate by reverse transcription of RNA intermediates from the positive-sense DNA contained in the virion. This is accomplished by a virus-coded reverse transcriptase (Seeger *et al.* 1986). The RNA serves as the template for a new (-) strand DNA. This in turn is transcribed by a DNA-dependent DNA polymerase to positive-stranded progeny DNA for incorporation into the new virions. It is currently assumed that replication requires conversion of the virion DNA into a double-stranded, covalently-closed circle (an episome), a reaction which is possibly mediated by the virion polymerase.

As shown in figure 1.5, HBV has four ORFs, designated S (for the three surface, or envelope, proteins); C (encoding the core proteins hepatitis B c and e antigens - HBcAg and HBeAg, respectively); P (which specifies the viral DNA polymerase); and X (transcriptional transactivator). HBV makes very efficient use of its small genome, with much overlap of the different ORFs. In addition, a single ORF is sometimes used to code for several proteins, through the deployment of different initiator codons within the same ORF. Thus, two translation start codons (AUG) are found in ORF C. Translation from the first gives rise to a long protein, of which the first 19 amino acids are a signal peptide, enabling secretion from the cell, after modification, as HBeAg. The core protein, HBcAg, is translated from the second AUG of this gene. In addition to B and T cell epitopes peculiar to HBcAg or HBeAg,

both peptides contain shared epitopes, a result of the large number of shared amino acids. For this reason, HBeAg has been postulated to play a role as a tolerogen in nenonates born to infected mothers, as well as an immunomodulator after infection in adults.

Figure 1.5: *Genetic map of the three HBs proteins (subtype ayw)*



Transcription start sites are indicated. Numbers refer to the amino acids of the primary translation products

Reproduction of Fig. 8.3, p.144, Chapter 8, Section 2, of "Viral Hepatitis: Scientific basis and clinical management," by A. J. Zuckerman & H.C. Thomas, eds. (1993), Churchill Livingstone., Edinburgh

The envelope gene contains three domains, each encoding a different protein (see figure 1.5). The small protein, or hepatitis B surface antigen (HBsAg), is encoded by the S domain and comprises 226 amino acids. A terminal extension of HBsAg forms the middle protein, which is HBsAg plus an additional 55 amino acids encoded by the pre-S2 domain. The largest hepatitis B surface protein (large protein), which is about 400 residues long, contains 3 domains - pre-S2 (encoding 108-119 amino acids); pre-S1; and S. All three proteins are translated from distinct mRNA molecules (Yamamoto *et al.* 1994). The pre-S1 domain is one of the most variable regions of the HBV genome. This may be a reflection of intensive selection by immune pressure. However, unlike the hypervariable regions of HCV or HIV, pre-S1 is not mutated within a chronically infected person.

1.4.3 Pathogenesis

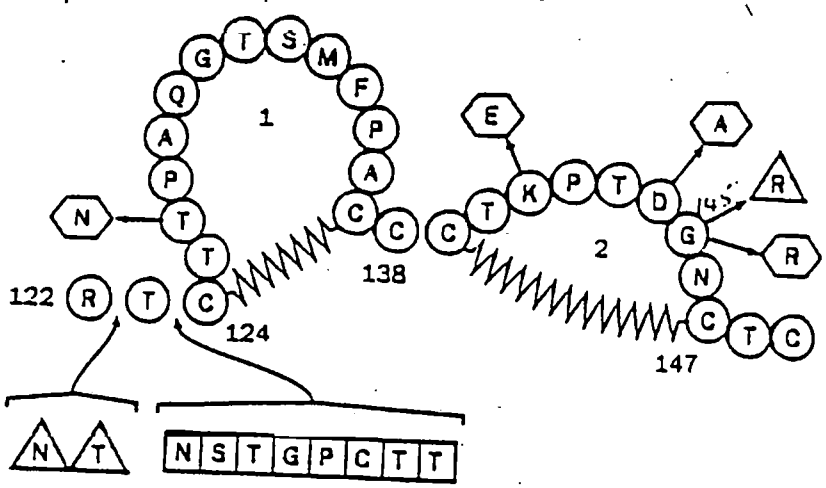
Although HBV is poorly cytotoxic *per se*, host defence mechanisms bring about the death of HBV protein-expressing cells. This, and the neutralisation of circulating virus, are the cause of acute and chronic hepatitis B. Proof of the immune-mediated nature of HBV-related liver damage lies in the observation that immunologically immature or deficient individuals become chronic carriers of the virus without developing typical hepatitis (Gerlich 1993). High viral titres (10^8 - 10^{10} particles/ml serum) are characteristic, and are accompanied by antigenaemia, in which the viral surface proteins are secreted by infected cells as particles of variable morphology and size. Most HBV-infected adults suffer symptomatic or asymptomatic hepatitis, and clear the virus. Less than 5% develop persistent infection with chronic hepatitis. During chronic infection, changes occur in all HBV genes, some of which result in evasion of the immune response. In a few individuals, an efficient cytotoxic immune response is followed directly by persistent infection with high-level viraemia and antigenaemia. The mechanism(s) whereby HBV suppresses or escapes immune pressure is (are) unknown. Surface protein variability does not play a major role, as is the case with HCV. Such variation is apparently distinct from the substitutions which define the subtypes of HBV. Although integration into the host genome is not essential for HBV replication, as is the case with orthoretroviruses, such an event often occurs, and may lead to disturbances of genetic elements of growth control. HBV therefore has oncogenic potential.

1.4.4 Genomic variability

Reverse transcription in viruses is a highly error-prone process, due to the lack of proofreading enzymes. Thanks to an intermediate reverse transcription step during replication, HBV evolves rapidly for a DNA virus, and is estimated to have a variability midway between that of DNA and RNA viruses (Wallace *et al.* 1994). One antigenic expression of this variability is seen in subtypes, which are defined by monospecific, polyclonal antisera against HBsAg. HBsAg has one common antigenic determinant, *a*, and one member of each of two pairs of mutually exclusive subdeterminants *d* or *y*, and *w* or *r*, to give 4 main antigenic subtypes: *adw*, *adr*, *ayw* and *ayr*. Additional subspecificities assigned

to the *w* determinant led to a 9-member classification: *adw2* and 4, *ayw*1-4 , *adrq+* and *q-*, and *ayr.* Figure 1.6 shows the predicted structure for the *a* determinant, which occupies the region from amino acids 124 to147, and is the main neutralising epitope cluster for HBV. Antibodies against *a* are a major component of the anti-HBs response following infection or immunization. In all serotypes, amino acid positions 139 - 147 are highly-conserved, and cyclical synthetic peptides of this nonapeptide have been shown to induce high-affinity antibody production (Karthigesu *et al.* 1994). As well as type-specific residues characterising particular sites (such as position 126, which is threonine in subdeterminant *w* and isoleucine for *r*), there are invariant residues. Four cysteines at positions 124, 137, 139 and 147, are predicted to form disulphide bonds (Cys 124-137 and 139-147) resulting in a double loop structure. The presence and conservation of these residues implies that strong secondary structural constraints operate in this region. The *a* determinant is found in all subtypes, so that HBV infection can be prevented by vaccination with a preparation of *a*-derived antigen (Waters *et al.* 1992).

Figure 1.6: *Proposed double loop structure of the HBsAg a determinant*



Circles represent normal residues; hexagons and triangles represent reported vaccine-induced mutants; squares depict insertion mutations Ref: Carman *et al.* (1995) - Lancet **245**:1406

1.4.5 HBV immunity: significance of the *a* determinant

The main marker antigens for HBV infection are produced sequentially as follows: Surface antigen (HBsAg), a non-infectious particle found in acute and chronic infection; core antigen

(HBeAg), whose presence implies infectivity; and core antigen (HBcAg), readily detectable only in hepatic nuclei, and not in blood. Patterns of infection are variable, and are influenced by age, sex and the state of the immune system. Typically, the incubation period for HBV is two to three months, and leads to a prodromal phase followed by overt jaundice, after which 90% of patients recover uneventfully. Antibodies appear in the order anti-HBc, anti-HBe and anti-HBs. Anti-HBs, the major protective antibody, appears only after HBsAg has vanished, and is therefore a reliable indicator of recovery and immunity to reinfection.

10% of HBV-infected individuals become chronic HBsAg carriers, in whom HBsAg, and often virions, circulate consistently in the blood. There are two kinds of HBsAg carrier state:-

- In chronic antigenaemia, the patient fails to form anti-HBs, and production of anti-HBe is delayed. In these "healthy" or "asymptomatic" carriers, HBsAg may persist in the blood for many years, but the patient is well, and does not infect others.
- Chronic active hepatitis B is distinguished by disease progression, as indicated by liver damage and persistently elevated serum levels of transaminases. Patients with this condition produce neither anti-HBs nor anti-HBe. Persistence of HBeAg, HBV DNA polymerase and Dane particles implies active viral replication, high infectivity, progressive liver damage and poor prognosis.

Testing for HBeAg , which may persist for years in chronically infected patients, is therefore of diagnostic value, as the antigen is a marker of viral replication (Martin *et al.* 1993).

In escape mutation, an antigen that has attracted immune attention is altered to become antigenically different from the initial epitope (Carman *et al.* 1993b). A virus bearing such an altered epitope may persist in the presence of an adequate immune response against the original epitope. The best known HBV mutants are those with mutations in the pre-core region for aborting HBeAg secretion. These mutants prevail as the host seroconverts to anti-HBe, because hepatocytes infected with HBV and surface expression of HBeAg are selectively eliminated. The link between severity of liver disease and such pre-core mutations is not well-defined - mutations are found both in patients with progressive disease and those with normal liver biochemistry (Hsu *et al.* 1995). It is unclear whether such substitutions predispose patients to progressive hepatitis or occur as a result of severe disease (Boner *et al.* 1995).

Variants of HBV with altered antigenicity of the envelope protein have been described. These are especially important because the *a* determinant bears potent HBV-specific B and T cell epitopes which form the basis of current vaccines and antibody-based assays for HBV. In a recent survey involving 1000 blood samples from individuals infected with HBV (Wallace *et al.* 1994), 0.8% were untypeable due to alterations in *a*. While representing a relatively low value, this figure is highly significant from clinical and epidemiological viewpoints. A variant of *a* with a substitution of Arg for Gly at amino acid position 145 was found in a patient who developed fulminant hepatitis after withdrawal of chemotherapy for lymphoma, despite previous vaccination (Carman *et al.* 1995). The patient had been anti-HBsAg seronegative in a widely used monoclonal antibody-based ELISA, but was found to be strongly seropositive by polyclonal assay as well as positive by PCR for HBV DNA. The Gly-Arg substitution, it should be noted, would significantly affect the hydropathy profile of the peptide, as well as its secondary and tertiary structure. Corroborating data showing that this variant evades the known protective anti-HBs response came from a study in which a vaccinated child born to an HBV-infected mother was found to harbour HBV with the point mutation predicting the Arg^{145} variant as the only mutation. This, and other changes involving single amino acid substitutions, as well as deletions from, or insertions into *a*, have been described (Waters *et al.* 1992; Karthigesu *et al.* 1994; Yamamoto *et al.* 1994).

Escape mutants, especially those affecting HBsAg and the *a* determinant in particular, pose a significant risk to the community because vaccines and immunoglobulin preparations are not effective in preventing infection with such mutant strains. Failure by current serological assays to detect HBsAg may lead to transmission through blood transfusions or organ transplants. Orthotopic liver transplantation (OLT) remains a useful method of treatment for end-stage liver disease caused by HBV, and is normally followed by administration of hepatitis B immunoglobulin (HBIG) in order to prevent recurrent infection. If OLT were performed on an individual carrying a mutant strain, HBIG preparations administered to prevent recurrent disease would be ineffective, and fulminant hepatitis could occur.

1.5 HEPATITIS C VIRUS: RNA AGENT OF HEPATITIS

1.5.1 Historical background

By the mid-1970s, routine testing for HBV carried out in blood banks world-wide had greatly reduced the incidence of hepatitis B transmission. Such testing also helped to underline the occurrence of other forms of hepatitis, known collectively as non-A, non-B hepatitis or NANBH (Bradley 1985; Tabor 1985). Two major forms of NANBH were prominent in the 1980s: those which were enterically and those which were parenterally transmitted (ET- and PT-NANBH, respectively). Physicochemical and histological studies showed significant differences between the two agents, such as sensitivity or resistance to chloroform, and type of liver lesion formed.

Similar to hepatitis B in its percutaneous route of transmission, the etiologic agent of PT-NANBH eluded detection for more than a decade, until the first definitive description by Choo *et al.* (1989), who based their work on pioneer experiments carried out with experimentally-infected chimpanzees. Previous work had shown a lack of abundant poly-A mRNA in the presumed viral agent (Linke *et al.* 1988). Underlying the investigation of Choo *et al.* was the assumption that the virus, although present in infectious serum, circulated in such low amounts that conventional immunological techniques had failed to detect it. Hence, as a viral source, plasma from a chimpanzee known to have a high degree of infectivity was used. After extensive ultracentrifugation, to ensure pelleting of even a small viral particle, nucleic acid was extracted from the pellet and completely denatured, and reverse transcription carried out on the assumption of either a DNA or RNA template, using random primers of reverse transcriptase. The resulting cDNA was cloned into a λgt11 library, and screened with presumed antisera. This led to the isolation of the first NANBH-associated clone (5-1-1), the product of which was used with first-generation ELISA tests for anti-HCV antibodies (Kuo *et al.* 1989). Most cases of post-transfusion NANBH were found to be associated with anti-c100-3 antibody, and this response was used to define infection with a new virus. The full genome of the etiologic agent, termed HCV, was subsequently cloned and characterized (Choo *et al.* 1991). Although other parenterally transmitted viral agents of hepatitis are now

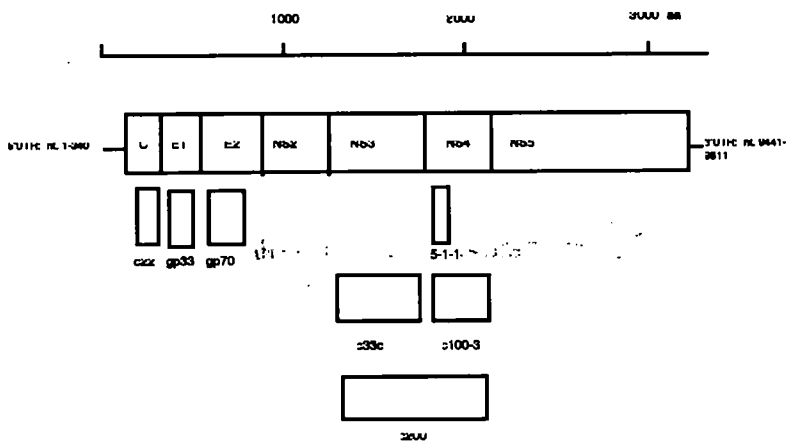
known to exist (Schlauder *et al.* 1995; Simons *et al.* 1995; Simons *et al.* 1995; Linnen *et al.* 1996), HCV still accounts for the majority of cases.

1.5.2 Epidemiology and risk factors for HCV infection

HCV is encountered worldwide, with relatively high prevalences in Japan, southern USA, the Mediterranean countries of Europe, Africa and the Middle East, where 0.1 - 1.5% of blood donors are anti-HCV positive. In the northern USA, Canada and northern Europe, prevalences are lower, at 0.01 - 0.05% (Van der Poel *et al.* 1991). Transmission is mainly by the parenteral route.

Screening assays for evidence of infection with HCV are usually based on commercial kits which detect antibodies against epitopes of the capsid (c22) and non-structural (c33, c100-3 and 5-1-1) proteins. The origins of the cloned antigens are depicted in Fig. 1.7.

Figure 1.7: *Schematic depiction of HCV genome showing origin of antigens used in commercial anti-HCV antibody detecting systems*



The implementation of anti-HCV antibody screening assays in blood banks around the world has led to a marked drop in the number of post transfusion hepatitis cases, with intravenous drug abuse (IVDA) now being the major risk factor for developing HCV infection. Transmission rates are especially high among haemophiliacs and dialysis patients (Allain *et al.* 1991; Ng *et al.* 1991; Brackmann *et al.* 1993; Lampertico *et al.* 1995). Other risk factors

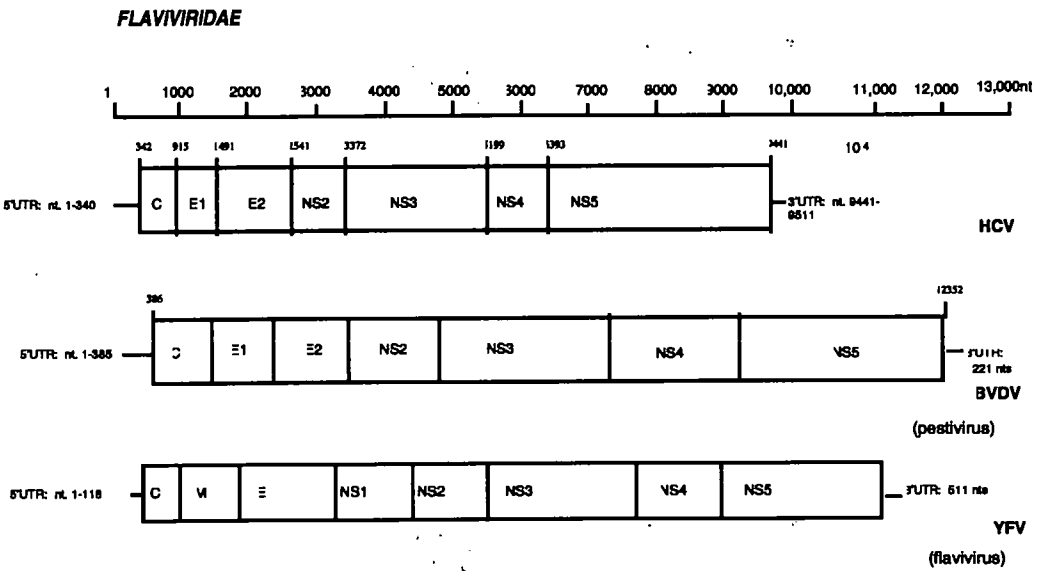
include tattooing (Ko *et al.* 1992), and needlestick injuries (Kiyosawa *et al.* 1991). Vertical and sexual transmission constitute uncommon but significant transmission routes (Alter *et al.* 1989; Melbye *et al.* 1990; Tedder *et al.* 1991; Inoue *et al.* 1992; Weiner *et al.* 1993). The demonstrated presence of HCV in body fluids such as saliva, urine, seminal fluid and ascites (Liou *et al.* 1992; Numata *et al.* 1993) indicates that these may play some role in nonparenteral HCV transmission. HCV is transmitted rapidly and efficiently to transplant recipients from infected donor organs (Wreghitt *et al.* 1994; Maple *et al.* 1994; Pereira *et al.* 1995).

1.5.3 Genomic organization

Alignment with sequences available from databases worldwide showed a significant resemblance between the newly-discovered HCV genome and those of flavi- and pestiviruses, as well as plant potyviruses and carmoviruses (Koonin 1991). The resemblance was greatest in areas encoding the enzymes of viral replication and expression, prompting the postulate of a common evolutionary ancestry for these positive-stranded RNA viruses (Koonin and Dolja 1993). According to this view, the genes for essential proteins (including RNA-dependent RNA polymerases, putative RNA helicases, chymotrypsin- and papain-like proteases, and methyltransferases) form a relatively stable core of "housekeeping" genes, interspersed with more variable domains. Limited but significant amino acid sequence homology was found in regions of HCV corresponding to the NS3 and NS5 proteins of flaviviruses (Miller and Purcell 1990). These included a nucleoside triphosphate (NTP) binding motif, as well as the putative (chymo)trypsin-like serine protease motif of flaviviruses and pestiviruses. In addition, the 3' end showed the Gly-Asp-Asp motif which is highly conserved among viral-encoded RNA-dependent RNA polymerases. The colinearity of these homologies, together with the overall resemblance of hydropathy profiles for the gene products, led to the proposal of a similar genomic organisation for HCV, flaviviruses and pestiviruses (Choo *et al.* 1991; Plagemann 1991; Takamizawa *et al.* 1991), with structural proteins encoded by the 5' terminus, and non-structural proteins by the 3' end of the genome. Fig 1.8 schematically

depicts HCV, bovine viral diarrhoea virus (BVDV, a pestivirus) and the representative flavivirus YFV (yellow fever virus).

Figure 1.8: *Genomic organization of HCV: Comparison with representative flavi- and pestiviral genomes*



UTR = untranslated region. Coding regions designated by letter corresponding to deduced protein encoded - M - matrix; E = envelope; C = core; NS = non-structural ; BVDV = bovine viral diarrhoea virus ; YFV = yellow fever virus

(Refs: Rice *et al.* 1985; Collet *et al.* 1988, 1989; Miller and Purcell 1990).

The 10 kb HCV genome contains a non-coding region (NCR) (or untranslated region, UTR) at the 5' end, which is highly conserved among all isolates. The following 9033 nucleotides constitute a major open reading frame (ORF) for the synthesis of a large polyprotein of 3010 or 3011 amino acids, which is further processed by cellular and viral genes to give the functional (three structural and four non-structural) proteins. The ORF is followed by a 3'-poly-A or poly-U NCR of variable length.

The 5'-UTR is a region of approximately 341 nucleotides (Takamizawa *et al.* 1991), which is highly conserved among variants of HCV worldwide, and bears some resemblance to a corresponding region in *Picornaviridae.* Picornavirus RNAs are uncapped messengers with unusually long 5' UTRs (610 - 1200 residues) which contain many silent AUG sequences - features which seem incompatible with efficient translation by the ribosome scanning

mechanism used in most eukaryotic cellular and viral mRNAs. Initiation of translation in picornaviruses involves binding of ribosomes to an internal sequence within the 5' UTR, called the internal ribosome entry site (IRES). In its possession of a fairly long 5' UTR harbouring three or four AUG sequences, HCV is more similar to picornaviruses than to *Flaviviridae*. Evidence that this domain in HCV constitutes an IRES was provided from studies using mono-

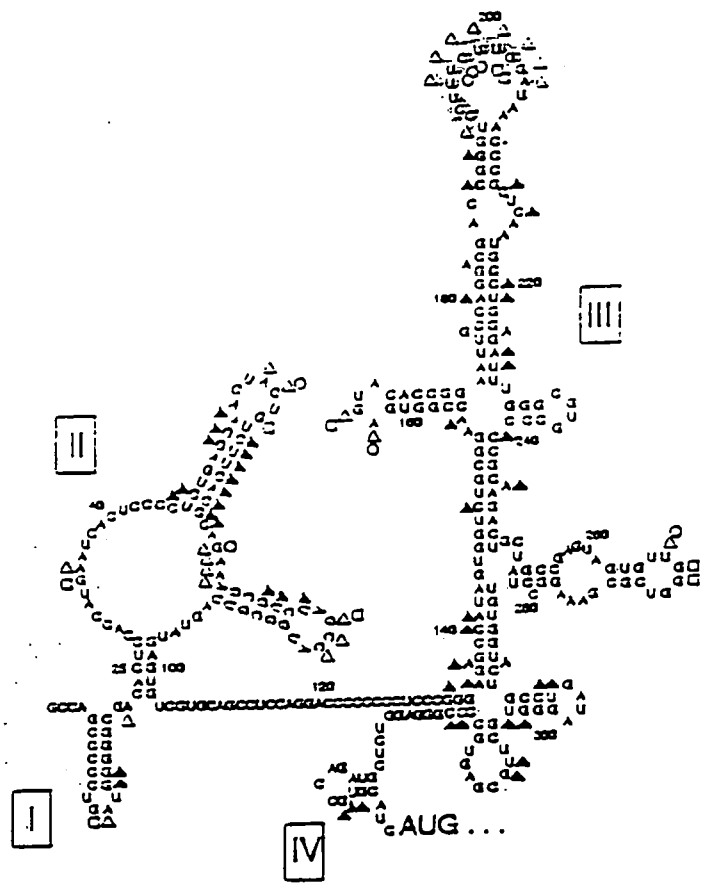


Fig 1.9: *Predicted secondary structure of the 5' untranslated region of HCV, showing four secondary structural domains (I-IV). Sites of nuclease cleavage are indicated thus: single-stranded specific T_1(square); T_2 (circle); S_1 (open triangle); and double-stranded specific V_1 (filled-in triangle)*

Ref: Brown *et al.* (1992) - Nucleic Acids Research **20**:5044, fig. 3

and dicistronic expression vectors, *in vitro* translation and deletion mutagenesis (Brown *et al.* 1992; TsukiyamaKohara *et al.* 1992; Wang *et al.* 1993). As well as demonstrating the 5' UTR involvement in translational control, these studies showed that the sequences immediately upstream of the initiator AUG are essential for IRES function during translation. Computer-assisted and biochemical analyses generated a thermodynamic model of the secondary structure of the HCV 5' UTR which was comparable to that found in picornaviruses. Depicted in figure 1.9, it composes a large conserved stem-loop structure (Brown *et al.* 1992).

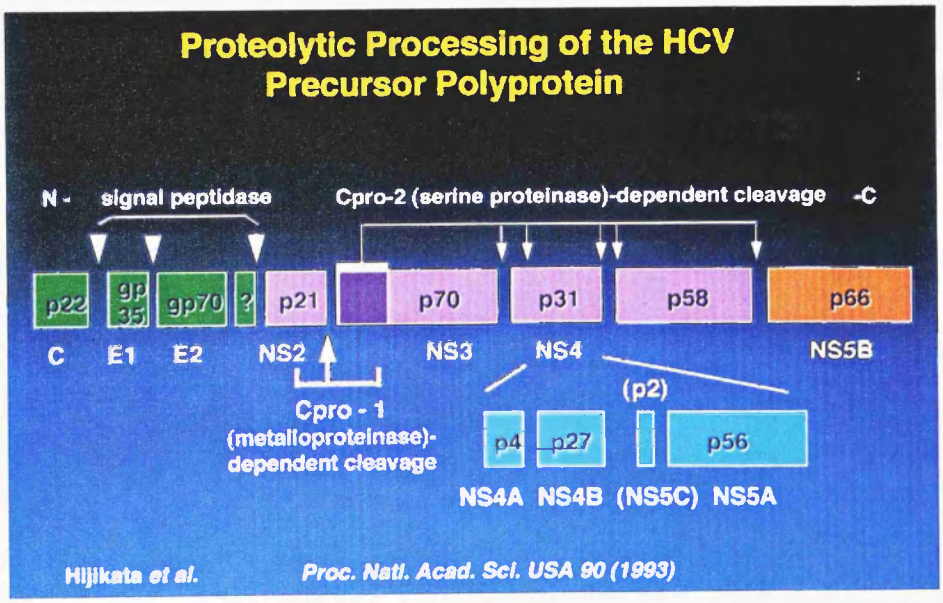
Other studies refuting the presence of an IRES in HCV indicate that the virus may employ an alternative, distinctive control strategy, involving the generation of subgenomic viral mRNA in infected cells (Yoo *et al.* 1992). The consensus remains, however, that translation strategies in HCV are different from those of other flaviviruses. This has prompted suggestions for the creation of a new virus family for classifying HCV.

The products of the entire coding region have been characterized by various groups (Eckart *et al.* 1993; Grakoui *et al.* 1993a; Ralston *et al.* 1993; Selby *et al.* 1993). Host signal peptidases are believed to effect cleavage between C and E1 and between E1 and E2, while processing of the NS proteins is by a virus-encoded protease. 9 distinct cleavage products are generated, viz:-

NH_2-C-E1-E2-NS2-NS3-NS4a-NS4b-NS5a-NS5b-COOH

As schematically depicted in figure 1.10, the basic nucleocapsid, or core, protein (p22) comprises 190 amino acids and is the first translated product. The predicted C gene product has a high (16-23.5%) Arg-Lys content, as do other flaviviral core proteins, and may be capable of binding to the RNA genome (Takamizawa *et al.* 1991; Okamoto *et al.* 1992). It is followed by regions encoding the hydrophobic envelope 1 protein gp33 (E1, 190 amino acids) and the second envelope protein gp70 (E2) of 350 amino acids. Although these are the most highly variable parts of the genome, they show a significant number of characteristics common to all variants, such as a constant number of cysteine residues, located at identical positions in the genome, and a remarkable conservation of potential N-linked glycosylation sites (Choo *et al.* 1991). A signal peptide motif (Perlman and Halvorson 1983; von Heijne 1985) precedes each envelope protein. NS2 (p23) is a metalloproteinase which mediates NS2-NS3 cleavage. NS3 (p72) has proteolytic activity responsible for some processing of the precursor polyprotein into mature products (Grakoui *et al.* 1993b). The function of NS4 is unknown, although it is known to be processed into NS4a (p10) and b (p27).

Figure 1.10: *HCV precursor polyprotein showing cleavage products*



Although the precise mechanism of HCV replication remains to be elucidated, it is significant that no DNA intermediates have been detected in serum or liver of infected individuals. This implies direct RNA to RNA replication, mediated by a virus-encoded RNA-dependent RNA polymerase. Antigenomic (-) RNA strands have been detected in liver (Takehara *et al.* 1992), plasma (Fong *et al.* 1991) and PBMC (Muller *et al.* 1993). Theoretical evidence for the polymerase role of NS5 has been reported (Takamizawa *et al.* 1991). Like NS4, NS5 is processed into a and b parts (p56 and p70). NS5b contains the GDD motif peculiar to RNA-dependent RNA polymerases of positive-strand RNA viruses and is assumed to encode the viral replicase (Koonin 1991). Taken together these data indicate that the HCV genome is replicated by a direct RNA-to-RNA mechanism, in which the RNA-dependent RNA polymerase encoded by NS5b catalyses the synthesis of the (-) strand on the (+) strand template, possibly mediated by host or viral factors which act as primers or initiation factors. The 3' UTR shows significant variations in both length and sequence in different reports, and may have a poly-U tail (Takamizawa *et al.* 1991).

Although overall nucleotide and amino acid sequence homologies between HCV and pestiviruses are few, a number of similarities point to a closer ancestral relationship to them than to flaviviruses. These include the similar size of the 5'-UTR (fig. 1.8), as well as the presence of nucleotide homology and of several short ORFs in this region. Also, both HCV putative envelope proteins resemble the pestivirus envelope glycoproteins. On the basis of the accumulated evidence, HCV is assumed to be an unusual virus that is most related to the pestiviruses.

1.5.4 Phylogeny and genomic diversity

RNA viruses are traditionally divided into positive-stranded viruses (in which the input genome is translated into protein); negative-stranded viruses (with genome complementary to the message sense); and double-stranded RNA viruses. As discussed in section 1.2.2, subdivisions into virus families are based on factors including virus structure, hosts and epidemiology. Comparative sequence analysis has revealed striking homologies among otherwise disparate groups, leading to the designation of taxonomic groupings such as the seven "superfamilies": the Sindbis-like, picornavirus-like, negative-stranded, double-stranded, flavivirus-like, coronaviridae and unassigned RNA viruses (Strauss *et al.* 1991). A hierarchical alignment of representative amino acid sequences from RNA-dependent RNA polymerases of all groups of positive-stranded RNA viruses was recently described (Koonin 1991). The exercise led to the delineation of three large supergroups of RNA viruses, designated I, II and III (Table 1.4).

Table 1.4: *Supergroups of RNA viruses*

I	II	III
Picorna-, noda-, como-, poty- and sobemoviruses	Carmo-, tombus-, pesti-, and flaviviruses; HCV; single-stranded bacteriophages	Tobamo-, tobra-, boralei-, tricorna-, potex-, toro- and rubiviruses; HEV

Unexpected clusters of sequences included the grouping of HCV and pestiviruses with carmoviruses and related plant viruses in supergroup II, and constitute strong evidence of a common ancestry. Extant viruses may have arisen from a single protovirus by a combination of linear divergence, genetic recombination and gene duplication.

HCV, like flaviviruses and pestiviruses, is a small, enveloped virus which, on the basis of the similarities described above, together with other, unique, features, has been classified as a separate genus within the family *Flaviviridae* (Heinz 1992). A comparison of full-length and partial sequences from around the world revealed the existence of different variants, which originally posed problems for detection of the virus. Experiments using PCR primer sets corresponding to different areas of the genome showed that primers derived from the non-coding region (NCR) gave the most reliable results (Bukh *et al.* 1992). The discovery of significant genomic diversity between different HCV isolates suggested the existence of distinct viral genotypes, and led to attempts to classify the virus. Comparisons of published sequences of HCV were carried out by several groups, leading to the identification of distinct viral 'types' which differed from each other by up to 33% over the entire viral genome (Choo *et al.* 1991; Chan *et al.* 1992; Okamoto *et al.* 1992; Bukh *et al.* 1993). Table 1.5 is a concordance of the major classification systems for HCV.

The Okamoto method for classifying HCV was based on differential PCR amplification using type-specific primers. In the two-stage PCR process, universal primers were first used to generate the template for second PCR. In the second stage, a universal sense primer was used in connection with a mixture of four antisense primers, each corresponding to a different HCV genotype (designated type I, II, III or IV). The generated DNA fragments were run on a gel and stained with ethidium bromide in order to visualise the type-specific PCR product, whose size was unique for each of the four genotypes. A fifth genotype (type V) was later identified (Okamoto *et al.* 1993). Primer sequences were based on a 341-nt sequence within the putative *C* gene. Nucleotide sequence identity within each genotype lay between 95.1% and 97.6%, a range significantly greater than that of 77.9 - 91% observed between any two types. This method was successfully used to type all of 256 serum samples tested, revealing

incidences of multiple infection by the presence, in each case, of multiple PCR bands (corresponding to different genotypes) within the same sample.

To identify evolutionary relationships between HCV variants, comparisons of nucleotide sequences from the 5' NCR were carried out on several isolates originating from different geographic areas (Chan *et al.* 1992). Phylogenetic analysis revealed the existence of three distinct groups of sequences, corresponding to designated HCV types 1, 2 and 3. As shown in figure 1.11, these phylogenetic groupings were maintained and further defined in the other regions analysed, NS3 and NS5.

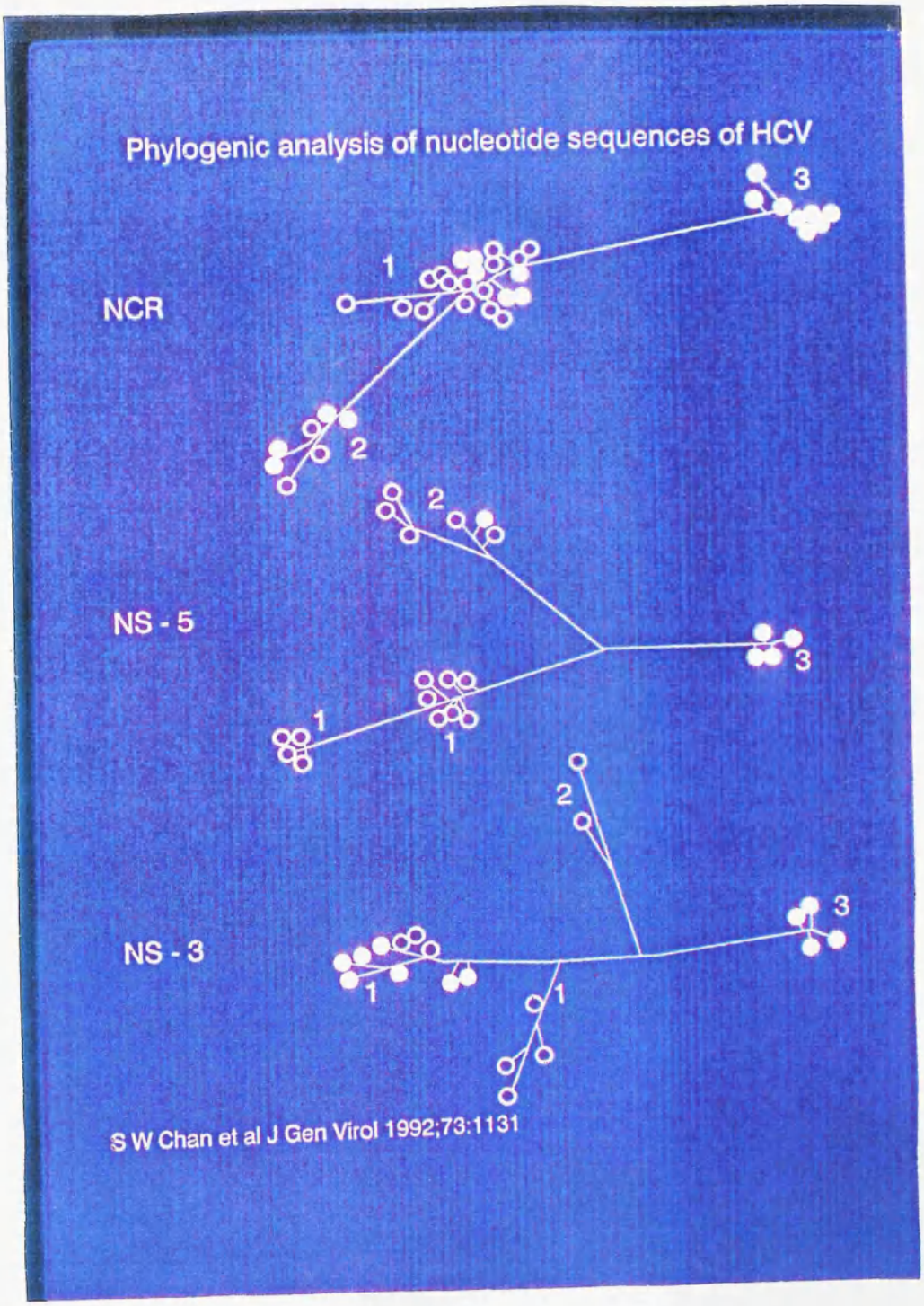
Table 1.5: *Concordance of major HCV classification schemes**

Proposed name	Published example	Cha	Simmonds	Enomoto	Okamoto	Tsukiyama -Kohara
1a	HCV-1, -H	I	1a	K-PT	I	NC**
1b	HCV-J, -BK	II	1b	K-1	II	I
1c	___	NC	NC	NC	NC	NC
2a	HC-J6	III	2a	K-2a	III	II
2b	HC-J8	III	2b	K-2b	IV	II
2c	___	III	NC	NC	NC	NC
3a	Ta, E-b1	IV	3	NC	V	NC
3b	Tb	IV	NC	NC	VI	NC
4a	___	NC	4	NC	NC	NC
5a	___	V	NC	NC	NC	NC
6a	___	NC	NC	NC	NC	NC

*Sequences from the following sources: Cha *et al.* (1992); Simmonds *et al.* (1993a); Enomoto *et al.* (1990); Okamoto *et al.* (1992); Tsukiyama-Kohara *et al.* (1991)

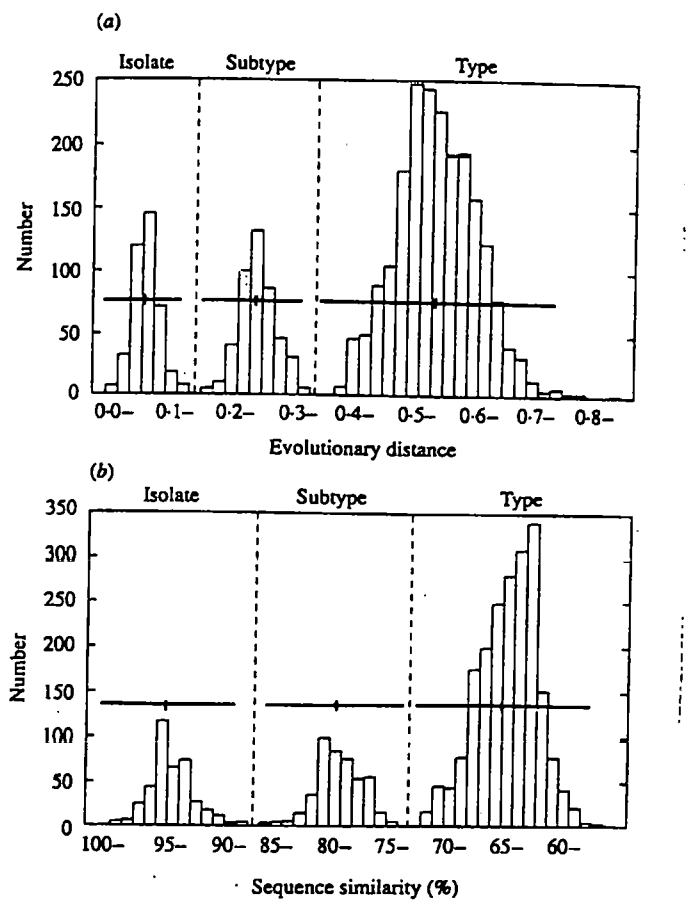
Adapted from Table 3, Simmonds et al. (1993a).

Figure 1.11: *Phylogenetic groupings (types) of HCV, based on analysis of the non-coding region (NCR), as well as nonstructural proteins 3 and 5 (NS3 and NS5) regions*



In later studies, NS5 was chosen for phylogenetic analysis due to the abundance of comparative sequence data for this region which was, in addition, sufficiently variable to allow differentiation between different HCV isolates (Simmonds *et al.* 1993a). Pairwise comparisons of 76 HCV variants allowed a detailed definition of the three overlapping groups into which the evolutionary distances were arranged (figure 1.12). The first ranged from 0.38 - 0.84 (mean 0.543); the second from 0.16-0.32 (mean 0.248); and the third 0-0.12 (mean 0.061). Of the 2850 pairwise comparisons made, no evolutionary distances of intermediate value between these distinct ranges were found, nor did the mean value of any distribution overlap with any other (figure 1.13).

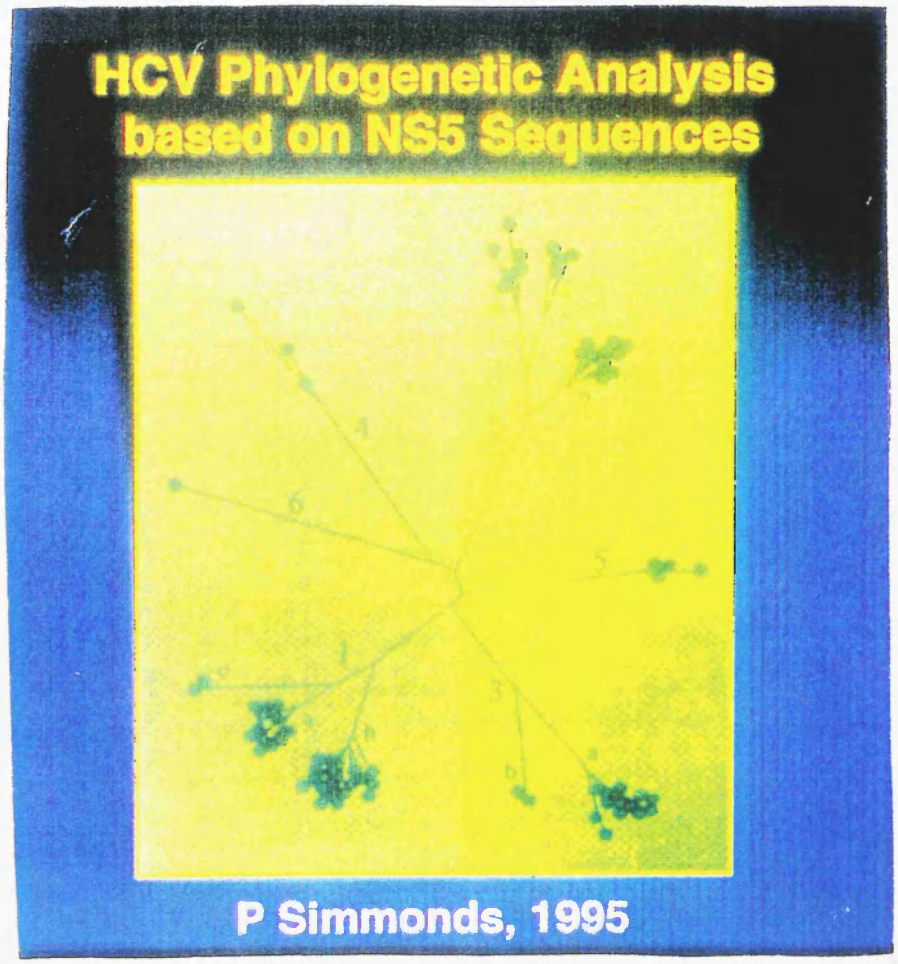
Figure 1.12: *Distribution of evolutionary distances (a) and percentage sequence similarities (b) upon pairwise comparison of 76 nucleotide sequences of HCV variants in the NS5 region. (a) No. of calculated evolutionary distance measurements recorded on y axis. (b) No. of observed sequence similarities recorded on y-axis. Mean ±3 S.D. for each distribution shown by horizontal bar*



Reproduction of Fig. 1, Simmonds *et al.* (1993a) Journal of General Virology **74**:2394.

Correspondingly, three levels of sequence diversity were observed in a phylogenetic tree of NS5 sequences. As depicted in figure 1.13, many types were found to contain groups of more closely related sequences, called subtypes and designated a, b, c etc. (Simmonds 1995). From these studies, it emerged that a 222-bp NS5 fragment was sufficient to distinguish among 6 equally divergent main groups of sequences, or types (designated 1 - 6).

Figure 1.13: *Evolutionary relationships between HCV variants (represented as numbered dots) based on NS5 sequence comparisons. 6 main groups (types) of sequence variants are shown (1-6). Types 1-3 each comprise groups of more closely related subtypes (designated a, b, c)*



Ref: P. Simmonds (1995) INTERaction **3(1)**:5

A typing method based on amino acid sequence homologies was proposed by Bukh *et al.* (1993). Using this method, it was possible to distinguish between at least 12 HCV types based on an 11-residue stretch of residues at the carboxyterminal E1 (amino acids 341 to 351). The practicability of the amino acid-based typing system led to its application in serological studies of the immune response during HCV infection using type-specific synthetic peptides or recombinant proteins. Results showed that HCV was classifiable into serological subtypes (Simmonds *et al.* 1993b; Orito *et al.* 1994).

In an extensive analysis of samples from 12 countries (Germany, Italy, Denmark, Sweden, USA, South Africa, Dominican Republic, Zaire, Hong Kong, Taiwan, India and Peru), HCV subtype 1b was found to be the most globally prevalent (Bukh *et al.* 1993). Certain (sub)types were nearly exclusive to particular geographic locations, such as the African types 4 and 5, and type 6 in Hong Kong. Interestingly, Denmark was found to have the highest number of subtypes, including the unique subtype 4d. Additional HCV groupings continue to accrue (such as types 7-9 in Vietnam), in step with ongoing attempts to clarify the significance of these variations (Bukh *et al.* 1995). Although two regions of hypervariability have been described for HCV, HVR2 is evident only in subtype 1b (Okamoto *et al.* 1993). The sequencing of additional HCV isolates from different parts of the world would outline the range of variability and substantiate classification based on a phylogenetic tree. Since HCV genotypes are associated with severity of liver disease and response to IFN, it is worth determing novel HCV types for correlation with liver disease and therapeutic efficacy, and for the development of serological assays and future vaccines.

The major mechanisms for creating sequence divergence in RNA viruses are reassortment (which only occurs in viruses with segmented genomes) and mutation. Mutations result from the fact that most RNA polymerases (replicases) lack proofreading activity, which leads to an inherent error frequency of 1 in 10^4 (Ogata *et al.* 1991). This, combined with the large number of replication cycles a virus can undergo in a year, and the occurrence of infection at high multiplicity (many viral particles/cell), makes for a very high rate of divergence. The rate of genomic sequence divergence in RNA viruses has been estimated at 0.03-2%/yr,

which approaches a million times the rate for eukaryotic DNA genomes (Holland *et al.* 1982). Although this figure may be overestimated, it is agreed that the overall mutation rate is so rapid that the sequence homologies found between, for example, plant and animal viruses can only be explained by common evolutionary pressures, which imply a common origin.

The issue of nomenclature of genetic variation is unresolved. The term "mutant" is used if a previous isolate from the same patient is shown to be different. For HBV, the term "variant" has been defined to mean a genome with altered biology (Carman *et al.* 1993a). The term "quasispecies" refers to the tendency of many viruses to circulate as a collection of distinct, albeit closely related, genomes, rather than a single population of homogenous sequences. All viruses consist of a mixture of viral genotypes, although the phenomenon is especially prominent in RNA viruses and retroviruses due to a lack of the proof-reading enzymes that assure fidelity of DNA replication (Holland *et al.* 1982). Mutation rates range from 10^{-6} to 10^{-3} point replacements, deletions or insertions per nucleotide per round of copying (Carman *et al.* 1993a,b). Each mutant progeny genotype in a heterogenous population can undergo any of 3 fates:-

- negative selection, a process in which mutants carrying changes lethal or deleterious to the virus are eliminated;
- random sampling of genomes with equal fitness for the environment. This results from the degeneracy of the amino acid codon which ensures that many individual point mutations have no phenotypic effect, or that changes do not affect RNA or protein secondary structure;
- positive selection by the host, in which mutants optimally fit for a new environment (such as escape mutants from immune pressure or mutants with a replicative advantage) dominate or replace the original genotype.

Random mutations, caused by errors of the viral polymerase , lead to the production of variants which are transmitted together and then replicate independently at different levels until one predominates. Positive selection, strongly represented by the host immune pressure in humans, is the major force which shapes the composition of a virus population within a

host. In immunocompetent hosts, the virus population usually consists of a single dominant sequence (master sequence), in coexistence with a spectrum of less well-represented sequences, designated minor variants. The development of antibodies against the major sequence allows one or more minor variants to selectively expand, a process that repeats itself as antibodies are sequentially produced against the sequentially emerging escape mutants. The quasispecies existence of viruses probably represents an evolutionary advantage, as it optimises the possibility of rapid selection of a variant with optimal fitness for any new environment, thus promoting viral persistence.

The HCV 5' non-coding and NS3 regions both have a confirmed "quasispecies" nature (Martell *et al.* 1992). The sequence heterogeneity rises to a peak in an 81-nucleotide hypervariable region of E2, designated HVR1 (Hohne *et al.* 1994; Kao *et al.* 1995). The composition of genomic variants in patients with chronic hepatitis C was found to change sequentially, either gradually or drastically (Honda *et al.* 1994; Kurosaki *et al.* 1994). The implication is that a population-based approach is vital in any studies to determine how the HCV genome evolves. For an accurate representation of the viral population in a host at any given time point, sequences from at least 10 clones must be determined for each region studied.

1.5.5 Evolution of genomic sequence during chronic infection

Early experiments aimed at characterising the HCV genome revealed that, although the nucleotide substitution rate of about 10^{-3}/site/year corresponded to that expected for an RNA virus, the changes were not uniformly distributed along the genome (Ogata *et al.* 1991). Rather, regions of high mutability were interspersed with relatively homogenous areas. Table 1.6 shows the observed nucleotide and amino acid sequence divergence between two HCV isolates, giving the values for different parts of the genome. These two isolates, which had been derived from the same chronically infected patient but separated by 13 years, showed an average nucleotide sequence divergence of 2.5%. However, values ranged from 0.7% in the NCR to 4.6% in E2. Furthermore, a 39-nt segment in the E2 was responsible for

30% of the observed changes. While changes in the other regions were predominantly (62%) in the third codon position, 100% of the changes in the highly variable region of E2 were in the first or second codon positions. Since nucleotide substitutions in most genes are known to occur preferentially at the third codon position, these results were strongly indicative of immune selection. Non-uniformity of nucleotide variation along the HCV genome was corroborated by others (Kremsdorf *et al.* 1991). In addition, the presence of two HVRs (HVR1 and HVR2) in the putative E2 was reported (Hijikata *et al.* 1991). A hypervariable domain, composed of 25 putative amino acids and designated region V, was reported for HCV, and mapped to the probable E1/E2 junction (Weiner *et al.* 1991). Region V, which included the 30 nucleotides of HVR1 - encoding 10 amino acids - as defined by Hijikata *et al.* (1991), represented 8% of the nucleotides sequenced, but accounted for up to 50% of the observed nucleotide changes, implying its being under immune selection or evolutionary/molecular constraints.

Table 1.6: *Differential variability of HCV genomic domains in two sequential isolates from a single patient*

Region	Changes (%)		% replacement mutations
	Nt	aa	
5' NCR	0.7	NA	
Core	1.4	1.4	33
E1	2.4	1.3	18
E2	4.6	7.2	52
NS2	3.1	1.8	20
NS3	1.9	1.2	20
NS5	1.9	1.5	25

Ref: Ogata *et al.*, (1991): Proc. Natl. Acad. Sci. **88**:3392-96

Comparisons of HCV genomes of different subtypes confirmed the even distribution of nucleotide sequence variability throughout most of the genome. As shown in Table 1.6,

exceptions are the core (invariably the most conserved) and non-coding regions, where sequence diversity is exceptionally low, and the highly variable envelope genes. The presence of a hypervariable region makes HCV unique among the *Flaviviridae*. The numerous features common to both region V3 of gp120 in the human immunodeficiency virus (the principal neutralising domain for HIV) and the HCV HVR1 have prompted attempts to clarify the clinical significance of the latter region by studying the nature and rate of sequential nucleotide changes therein.

It appears that the HCV genome remains remarkably stable in the period immediately following transmission. Vertical transmission studies reported a child with a homogenous population born to a mother with multiple HVR1 variants (Weiner *et al.* 1993), with similar results obtained from experimental transmission of HCV to chimpanzees (Van Doorn *et al.* 1994). When HCV sequence analysis was carried out on chronically-infected patients, however, serial samples were shown to undergo dramatic changes of nucleotide sequence (Kurosaki *et al.* 1993). The observed strong preference for nonsynonymous mutations constituted indirect evidence of *in vivo* positive immune selection. The results of a 28-patient study reported in 1994 indicated that the degree of intrapatient HVR1 heterogeneity increased significantly with progression of liver disease (Honda *et al.* 1994). In this report, a total of 280 clones (10 per patient) were sequenced in the region corresponding to the HVR1, now defined as an 81-nt region encoding putatively 27 amino acids from positions 384 - 410, which includes the 25 amino acids of region V. Multivariate analysis showed that histological diagnosis was the strongest indicative factor affecting nucleotide diversity, so that in acute hepatitits few variants were detected, while in patients with cirrhosis and hepatocellular carcinoma, all 10 clones were different, with no master sequence.

1.5.6 The E1 and E2 domains

The enormous diversity of RNA viruses probably leads to mutations introduced uniformly along the genome. However, mutations in regions where a high degree of conservation is absolutely required would give rise to defective viruses, which would rapidly disappear. The

observation of a highly variable envelope region, for example, may indicate that envelope proteins are more tolerant of amino acid mutations than other proteins, and therefore persist. The envelope proteins of viruses play a major role in stimulating protective immunity against them, as discussed for HBV in section 1.4. In order to investigate possible targets for the host immune response, the envelope glycoproteins of HCV have been extensively studied. Both E1 and E2 are heavily glycosylated proteins (Grakoui *et al.* 1993). Cleavage positions for E1 and E2 have been mapped at approximately amino acid positions 190 and 380, respectively (Matsuura *et al.* 1992).

E1: The putative product of the HCV *E1* gene is a glycosylated protein comprising amino acid positions 192 to 383 of the initial polyprotein. 192 residues in length, it contains a moderately variable domain between amino acid positions 215 and 255 (Weiner *et al.* 1991). The E1 protein has a size of 31 kDa, and at least 5 potential acceptor sites for N-linked glycosylation, represented by the sequence Asn-X-Thr/Ser, where "X" = any residue (Grakoui *et al.* 1993). Underlying the apparent diversity of this glycoprotein is the occurrence of certain invariant features, including a 15-residue stretch from amino acid 315 to 329, depicted below:-

GHRMAWDMMMNWSPT

This stretch of amino acids is absolutely conserved in all HCV isolates, except for the substitution of Leu for Met at position 324 in HCV type 2 (see section 1.5.4) (Okamoto *et al.* 1992). Potential glycosylation sites are well-conserved in E1. Disulphide bond formation plays an essential role in determining the three-dimensional structure of E1, however variable the intervening amino sequences. This is reflected in the invariant number and positions of its 8 Cys residues (Okamoto *et al.* 1992). Processed E1 protein expressed in mammalian and insect cells was examined for antigenicity in a Japanese study (Matsuura *et al.* 1992). In a population of 360 samples from patients with hepatitis C of varying severity, the prevalence of anti-E1 antibody was generally low at about 20% (compared with 90% for anti-core protein antibody). The reason for this discrepancy could be the low immunogenicity of E1, or hypervariability in this region.

Proteolytic processing of the E1, E2 and core proteins is dependent only on host signal peptidase (Ralston *et al.* 1993). A signal sequence is a labile peptide, 13 to 36 amino acids long, found at the amino terminus of nearly every presecretory polypeptide, which is cleaved from the mature protein after initiating transport. A target of signal peptidase, it is invariably composed of three structurally dissimilar regions: a positively-charged amino-terminal domain (*n*); a central hydrophobic core (*h*); and a more polar carboxyl terminus (*c*) that apparently defines the cleavage site. The minimal length of a signal peptide substrate of a eukaryotic enzyme has been defined as 13 amino acids, comprising *n*, *h* and *c* regions of 1, 7 and 5 amino acids respectively (von Heijne 1985).

A comparison of flaviviral, pestiviral and HCV hydropathy plots indicated the likely presence of a signalase cleavage site in the region of HCV encoding amino acids 371 - 377 (Weiner *et al.* 1991). Named region SP, this could represent the hydrophobic core *h* of the signal peptide. The same study revealed the presence of a hydrophobic domain between amino acids 328 and 368. The residues involved are:

PTTALVMAQLLRIPQAILDMIAGAHWGVLAGIAYFSMVGN

in the prototype HCV isolate HCV-1 (Choo *et al.* 1991). This region, which is located directly upstream of SP, could serve as a membrane anchor sequence, as hypothesised for flaviviruses, and lends support to the postulated function of region SP. Alanine, the most abundant protein associated with signal peptide cleavage (Perlman and Halvorson 1983), is found five residues downstream of region SP, and a (positively-charged) Lys residue precedes it. Together, these data indicate that the signal sequence for E2 cleavage is as follows:-

KVLVVLLLFAGVDA!ET,

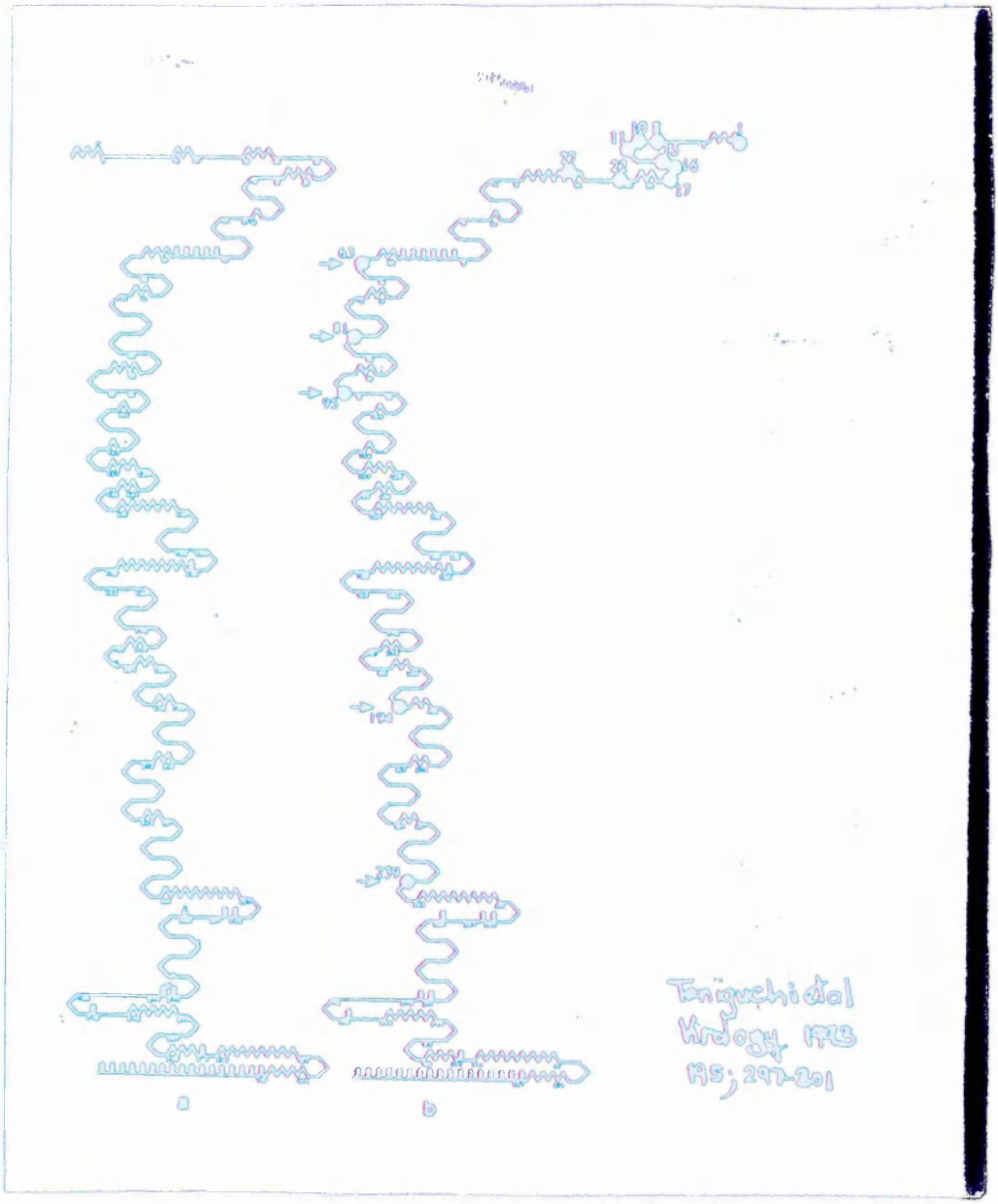
running between residues 370 and 383, where region SP is emboldened, and "!" indicates the signal peptidase cleavage site. The residues E (Glu) and T (Thr) are the first two amino acids of a known hypervariable region (HVR1, discussed below). The study of Weiner *et al.* was useful in pinpointing this site of cleavage between E1 and E2. It should be noted that

cleavage at the position indicated would place HVR1 at the extreme amino terminus of the mature E2 protein (see fig. 1.14).

E2: The *E2* gene product is a 70 kDa glycoprotein comprising about 350 amino acids (Grakoui *et al.* 1993). Formerly designated E2/NS1 to reflect the role of analogous proteins in flaviviruses and pestiviruses, its heavy glycosylation, together with the fact that it is not secreted into the culture medium by transfected mammalian cells, constitute strong evidence that it is a second envelope protein. Its 11 potential N-linked glycosylation sites are well-conserved, and represent many more than the number of such sites in its putative analogs - a maximum of 5 in the flaviviral NS1, or 6 in pestiviruses (Spaete *et al.* 1992). Other fixed features indicating the presence of conformational restraints include the possession of exactly 26 Cys residues at identical positions in different isolates, and an abundance of Gly and Pro residues. Thus E2, as well as E1, appears to have a complex three-dimensional structure. Four hydrophobic sections between residues 715 and 779 are believed to compose a single, membrane-spanning domain. In the proposed topology of Spaete *et al.* for the E2 protein, this would place the glycoprotein in a conventional orientation with the N-terminus out of and the C-terminus within the cell membrane. Thus, the N-terminus would be subject to immune surveillance.

Hypervariable regions of E2: Two hypervariable regions (HVR1 and -2) have been identified in E2, which account for its being the most diverse domain of the HCV polypeptide. Most of the observed diversity is contributed by HVR1, which encodes a 27-residue polypeptide occupying amino acid positions 384 to 410 of the polypeptide. HVR1 has been shown to be a target of the immune response which stimulates the production of neutralizing antibodies (Shimizu *et al.* 1994; Zibert *et al.* 1995). Although neutralizing, the antibodies do not prevent chronic infection in most cases because of the rapid changes which occur progressively, behind which the isolate-specific antibodies lag. Amino acid alterations in HVR1 occur sequentially during chronic hepatitis at a rate of 0.5 - 1.7 amino acids/month (Kato *et al.* 1994).

Figure 1.12: *Predicted secondary structure of E2 protein*



White dots indicate sites of amino acid substitutions, deduced from nucleotide sequences of HCV isolates from a chimpanzee host. These were obtained during the acute phase of infection (a) and 8 years later (b). DDBJ/EMBL/Genbank Accession Nos. are D01217 and D10750, respectively. Only amino acid changes in HVR1 (positions 1-27 above) produced an alteration in local secondary structure

Ref: Taniguchi *et al.* (1993) - Virology **195**:298

The HVR2 domain is a nonapeptide or heptapeptide, depending on HCV type, whose product occupies amino acid positions 474 - 482 of the type 2 HCV isolate HC-J8 (Okamoto *et al.* 1992). HVR2 apparently occurs in only one subtype (1b) of HCV, and its product does not appear to be a target of the immune response (Okamoto *et al.* 1993).

The result of computer-generated secondary structural predictions for the E2 protein is shown schematically in figure 1.14). The amino acid sequence for E2 was determined for two HCV genome isolates obtained from the same chronically infected host, one during acute infection and the other 8.2 years later (Taniguchi *et al.* 1993). The 27 amino acids of HVR1 constitute less than 0.1% of E2. However, the changes in amino acid sequence between the two isolates were concentrated in this region, which contained 9 of the 14 observed substitutions. The secondary structure of the HVR1 domain was significantly different at the second time point, while the bulk of E2, that part outside HVR1, showed no alteration in predicted secondary structure despite 5 amino acid changes. These findings helped to reconcile the apparent constraints on secondary structure of the envelope glycoproteins with the ease of escape from neutralising antibody that makes hepatitis C infection predominantly chronic in nature. HVR1 appears to be the single "loophole" in the highly conserved architecture of E2. Its presumed orientation outside the cell membrane would render it easily accessible to neutralising antibodies, while its occurrence at one terminus of the protein means that many changes may be tolerated without affecting overall secondary structure. Antigenic shift during the course of chronic infection leading to escape from neutralising antibodies is one of the most probable mechanisms of HCV persistence.

E2 protein could be the structural analogue of hog cholera virus (HoCV) gp55, which has been shown to elicit protective antibodies against hog cholera when vaccinated into swine (Hulst *et al.* 1993). In one experiment designed to check this theory, truncated and full-length versions of the E2 domain were stably expressed into CHO cell lines (Spaete *et al.* 1992). While the full-length form appeared to be exclusively intracellular, the C-terminally-truncated forms were detectable in extracellular media as fully processed glycoproteins containing terminal

sialic acid additions. These truncated proteins could be biologically relevant targets of the immune response, and therefore constitute potential subunit vaccine candidates.

Other regions of E2 have been investigated for immunogenicity. The antibody response to structural and nonstructural proteins was studied in immunocompetent patients, and compared to the response with immunosuppressed patients, all HCV-infected (Lok *et al.* 1993). It was found that most immunocompetent hosts with chronic hepatitis C were highly reactive to all antigens tested (derived from the core, E1, E2, NS3 and NS5 regions). Reactivity was much decreased in immunosuppressed patients, but this effect was less pronounced with core- and envelope-directed antibodies. This study highlighted the advisability of incorporating assays for additional antigens, including E1 and E2, into serological assays, at least until more sensitive PCR-based tests become commercially available.

In a study to further characterize the antigenic structure of E2, 70 synthetic, overlapping peptides spanning E2 were used in serological assays for specific antibodies. Two major antigenic areas were found, occupying amino positions 484-499 and 554 - 569 (Zhang *et al.* 1994). The sequences, which were almost totally conserved among the isolates tested, were found in 50% of 131 patients with acute, past or chronic HCV infection. However, there was little evidence suggesting that antibodies directed against these regions are neutralizing.

Experiments using purified E1-E2 complexes revealed that they are recognised at high frequency by HCV-positive human sera (Ralston *et al.* 1993). The antigenicity of E1 appeared to be affected by co-expression of E2 in *cis*. The molecular basis of E1-E2 complex formation was unclear in these studies. Although Grakoui *et al.* (1993a) had earlier reported the occurrence of intermolecular disulphide linkages, this finding was not corroborated by the results of Ralston *et al.* The requirement for E1-E2 co-expression for high antigenicity provided a possible explanation of the weak immune response to recombinant E1 expressed from constructs lacking E2, and implied that both E1 and E2 might be necessary components of an HCV vaccine. Indeed, further studies showed that chimpanzees vaccinated with

vaccinia-expressed E1 and E2 were completely protected from infection after intravenous challenge with the homologous HCV-1 isolate (Choo *et al.* 1994). The combined results form an encouraging basis for the development of anti-HCV vaccines which are universally applicable.

1.5.7 HCV immunity: role of HVR1

Two peculiarities which distinguish HCV from members of other *Flaviviridae* genera are the frequent chronicity of HCV infection, and the extreme heterogeneity of the N-terminus of the E2 polypeptide. The idea of a correlation between these two facts was pursued through investigations into the possible immunogenicity of the hypervariable region. In accordance with the general rule that linear epitopes are associated with less-structured regions of proteins, such as the ends, or surface loops, secondary structure analysis of the N-terminal E2 HVR was carried out. Resulting data indicated a relatively unstructured nature of this region, which had <50% probability of α-helix, β-turn or β-sheet character (Weiner *et al.* 1992). The lack of secondary structural motifs in the hypervariable region, designated region V by Weiner *et al.*, implied a tolerance for sequence changes compatible with its putative role in escape from immune pressure, and antibody-epitope binding studies showed the presence of isolate-specific linear neutralising epitopes in region V (Weiner *et al.* 1992).

Further investigations of the antigenic potential of predicted HVR1 peptides obtained from serial samples revealed a successive appearance of mutants with antigenically distinct amino acid sequences within this domain (Taniguchi *et al.* 1993). Each amino acid substitution was associated with alteration of the predicted local secondary structure of the epitope region, implying a high degree of structural flexibility and antigenic variability. These properties, which would facilitate ease of escape from host immunity, may well account for the high persistence rates of HCV.

In an assay for HCV-specific neutralising antibodies, cultured cells were incubated with either high- or low-infectivity titre samples (HITS or LITS) as an HCV source (Shimizu *et al.* 1994). The LITS, known to be antibody-complexed, were unable to sustain intracellular viral

replication, implying that the complexed immunoglobulins were neutralising antibodies. Although the particular epitopes involved were not identified, it seemed reasonable to assume that they would be located in the viral envelope, as is the case with many other viruses, including *Flaviviridae*. Both the pestiviral E2 and flaviviral NS1 proteins elicit protective antibodies against the respective viruses when used as vaccines (Weiner *et al.* 1992), and pigs vaccinated with the purified hog cholera virus (HoCV) E1 product were protected against hog cholera (Hulst *et al.* 1993). With HCV, *in vitro* experiments, in which humoral responses to HVR1-derived peptides were evaluated with expression vectors producing such peptides, pointed to a key role of the HVR1 in the immune response. Immunoprecipitation occurred upon the addition to this system of homologous, but not heterologous, sera. These findings led to the identification of immunological epitopes specific for the homologous viral isolates (Kato *et al.* 1994). Two distinct, overlapping epitopes of at least 11 peptides each were identified, running from amino acid positions 11 - 21 and 14 - 24 of HVR1. Variants altered within these epitopes could escape recognition by pre-existing antibodies, implying that HVR1 is the major site affecting genetic drift in HCV. Corroborating evidence came from physicochemical studies which revealed the presence of two populations of HCV particles (high-density and low-density) in sera of chronically-infected HCV hosts, the high-density component deriving from association of the virion with immunoglobulin. Sequence analysis showed that only the low-density particles accumulated base substitutions within HVR1, implying that such changes can lead to accumulation of immunoglobulin-free viral particles (Choo *et al.* 1995). In a further experiment to test the hypothesis that these antibodies recognised specifically HVR1, 13 clones derived from an HCV-contaminated anti-D preparation were sequenced (Zibert *et al.* 1995). The expression products of the four identified HVR1 variants were used in an ELISA assay for anti-HVR1 antibodies. Such antibodies were found in 67% of serum samples from 30 recipients of the anti-D source, versus in 0% of 60 negative controls.

A major difference between HCV and other *Flaviviridae* is the apparent lack of protection from infection by antibodies to envelope proteins. Previous HCV infection did not protect

chimpanzees experimentally inoculated with the virus from hepatitis - seroconversion and liver alterations developed regardless of whether the reinfecting strain was homologous or heterologous to that of the original inoculum (Farci *et al.* 1992; Prince *et al.* 1992). A similar phenomenon has been described in humans: A patient became superinfected with a type III (subtype 2a) HCV strain despite pre-existing subtype 1b infection (Kao *et al.* 1993).

1.5.8 Chronic hepatitis C: Mechanism of pathogenesis

Little is known about the pathogenesis of HCV, a virus which is strongly associated with liver injury in infected humans. Studies of this phenomenon are complicated by the absence of an efficient *in vitro* culture system, or convenient animal model. Early indications of a positive correlation between HCV viraemia levels and abnormal liver histology were confuted by clear evidence of persistently high hepatitis C viraemia in the absence of liver disease (Brillanti *et al.* 1993; Navas *et al.* 1994; Barrera *et al.* 1995). These findings imply that a direct cytopathic effect is not the major cause of liver damage in hepatitis C, a view supported by the fact that the virus lacks oncogenes and does not replicate via a DNA intermediate.

Studies on congenitally immunocompromised patients have provided further clues to the nature of HCV pathogenesis. Patients with primary hypogammaglobulinaemia who were treated with contaminated immunoglobulin rapidly became infected with HCV, and experienced a severe and rapidly progressive course of hepatitis which was poorly responsive to IFN (Bjoro *et al.* 1994). Sequence analysis of the HVR1 in an agammaglobulinaemic patient revealed absolute homology of consensus sequences at three different time points spanning a two and a half year period (Kumar *et al.* 1994). These observations imply that a humoral immune response against HVR1 is not a necessary accompaniment to HCV-related disease. However, the host immune system may play a role in HCV-related liver injury via its cellular branch. Possible mechanisms include T-cell-mediated cytotoxicity; and the release of inflammatory cytokines, mediated by HCV within leukocytes.

Investigations into the nature of cellular immunity in HCV infection, which revealed the existence of HCV-specific cytotoxic T lymphocytes (CTLs), have answered some of the questions concerning the mechanism of HCV persistence and pathogenesis. Human histocompatibility antigen (HLA)-restricted CTL recognising epitopes from various parts of the HCV genome have been isolated from both chronic HCV patients and asymptomatic anti-HCV positive subjects (Ferrari *et al.* 1994). The highly-specific HLA class II-restricted T cell responses described in the study of Ferrari *et al.* were directed against the core protein, which appears to be the most potent T cell antigen for both chronic HCV patients and asymptomatic, anti-HCV-positive subjects. In another study, HLA class I-restricted $CD8^+$ CTLs were isolated from liver-infiltrating lymphocytes of persons with chronic HCV hepatitis (Koziel *et al.* 1993). A key aspect of this finding was the confinement of the CTL to the sites of tissue injury, which supports the hypothesis of CTL-mediated liver injury. The complexity of the subject is highlighted by the fact that HCV-specific, HLA class I-restricted $CD8^+$ cells could also be isolated from lymphocytes in the peripheral blood circulation (Kita *et al.* 1993).

It is clear from the above that both HLA class I- and class II-restricted T cell responses to HCV occur, directed against antigens from various parts of the genome. The fact that significant variation existed in some of the epitopes described, particularly those within E2 and NS2 (Koziel *et al.* 1993), implicates mutations within CTL epitopes as a mechanism for viral persistence through escape from the cell-mediated immune response. Antagonism, a phenomenon whereby mutations within T cell epitopes prevent binding or recognition of the original peptide, may also play a role. Natural amino acid substitutions within the HBV core region in HBV-infected patients were found to affect CTL function through the antagonistic effect of variant peptide-HLA complexes on TCR recognition of the wild-type peptide (Bertoletti *et al.* 1994). Such a mechanism could operate in HCV which, like HBV and HIV, displays uncommonly high mutation rates, and could prove problematic in developing universally applicable therapeutic strategies. One encouraging finding was the widespread occurrence of HCV core-derived CTL epitopes in infected patients, common to all studies described in the above paragraph. In addition, a highly conserved antigenic site in HCV core

has been described (Shirai *et al.* 1994). Recognised by both human and murine CTLs in association with the highly prevalent HLA class I molecule A2, this conserved nonapeptide may be a valuable component of HCV vaccines against a broad range of HCV isolates.

Immunosuppression, of congenital or therapeutic origin, is associated with higher viral titres and more severe liver injury than usual. These features are especially severe in the post-transplantation setting when immunosuppression is increased for the treatment of rejection (Sheiner *et al.* 1995). The source of graft infection in these patients is unclear. While virions remaining in the blood at the time of transplantation may play a role, the finding of HCV within peripheral blood mononuclear cells (Bouffard *et al.* 1992; Moldvay *et al.* 1994; Navas *et al.* 1994) suggest these as a possible reservoir. There is increasing evidence for differential pathogenicities of the different HCV subtypes, with subtype 1b being particularly virulent (Hino *et al.* 1994; Pozzato *et al.* 1994; Simmonds 1995). However, this may merely reflect the time of infection. HCV infection has been associated with autoimmune disease in certain patient populations (Garson *et al.* 1991; Agnello *et al.* 1992; Bloch 1992; Michel *et al.* 1992; Agnello 1995). A strong correlation has also been found between HCV infection and hepatocellular carcinoma (Hasan *et al.* 1990).

1.5.9 Therapeutic strategies for chronic hepatitis C

At present, the only effective method for treating hepatitis C is interferon therapy. The interferons (IFN) are naturally occurring peptides with antiviral, antitumour and immunomodulatory properties which have significant biochemical and histological effects when used to treat viral hepatitis (Shindo *et al.* 1991; Lucey 1995). In HCV treatment, although IFN was able to restore normal liver biochemistry and reduce or eliminate HCV viraemia during therapy in most cases, it had a high relapse rate, with less than 20% of treated patients achieving a sustained response (Hoofnagle *et al.* 1986; Garson *et al.* 1992; Diamond *et al.* 1994). Also, viraemia levels and IFN response may vary among HCV types due to differential replication or differential host susceptibilities (Yoshioka *et al.* 1992). Intraisolate HCV genomic heterogeneity was associated with persistent infection and resistance to IFN,

while prognosticators for a sustained response included absence of subtype 1b, low nucleotide sequence diversity, youth, female sex and low-level viraemia (Enomoto *et al.* 1994; Hino *et al.* 1994; Kanazawa *et al.* 1994; Koizumi *et al.* 1995). Rates of sustained response may be increased by increasing the period of therapy (Shindo *et al.* 1991).

Ribavirin (1-β-D-ribofuranosyl-1,2,4-triazole-3-carboxamide) is a non-IFN-inducing guanosine analogue with a broad spectrum of activity against DNA and RNA viruses, including flaviviruses (Reichard *et al.* 1991; Sherlock 1995). It acts at the level of translation by disrupting cap formation. The first drug to offer potentially effective oral treatment for chronic hepatitis C, it induces a clearcut biochemical response, but has no effect on viral RNA, and biochemical relapse invariably occurs after ceasing therapy (Reichard *et al.* 1991; Camps *et al.* 1993).

Reports at a recent conference indicated that a combination of IFN and ribavirin was much more effective than IFN monotherapy (Alberti and Chemello 1995; Brillanti 1995; Weiland *et al.* 1995). This effect was apparently synergistic, and was especially remarkable among relapsed responders to IFN monotherapy. In a pilot study of combination therapy in immunosuppressed graft recipients, a sustained response was observed in 12 of 14 liver transplant recipients with post-transplantation HCV reinfection. These results were particularly encouraging because the post-transplantation situation combines the worst prognosticators for successful IFN monotherapy: high viral titres of HCV subtype 1b in immunosuppressed patients (Bizollon *et al.* 1995; Main 1995). However, the exact nature of the relationship between viral titres and severity of liver disease is not clearcut.

1.5.10 Orthotopic liver transplantation as management strategy

A transplant is tissue (an organ or part thereof) which is removed from its original site and transferred to a new location on the same (autograft) or another person (allograft). A graft transfer between identical twins is an isograft, and a heterograft, or xenograft, is transferred from one nonhuman animal to a human. Autografts and isografts are accepted

indefinitely by recipients, allografts may be successful when special efforts are made to prevent rejection, and xenografts are usually rejected. In an orthotopic transplant, the organ is grafted into its usual position, in contrast to heterotopic transplantation, where the organ occupies an auxiiliary position.

Rapid advances in the treatment of liver disease have been made possible by the development of transplantation procedures. Orthotopic liver transplantation (OLT), pioneered by Thomas Starzl in the mid-50s in the USA, and later by a University of Cambridge/Kings College London consortium, is now used as a last resort in most cases of lethal hepatic disease (Starzl *et al.* 1989a; Calne 1993). The most common diseases treated are chronic active hepatitis (CAH), primary biliary cirrhosis (PBC), alcoholic cirrhosis and inborn metabolic errors (IME).

Despite reports of permanent or prolonged liver graft acceptance by untreated pigs or minimally suppressed dogs, there is evidence of graft rejection in at least 70% of human cases (Starzl *et al.* 1989b). Rejection is prevented by the administration of immunosuppressive drugs, of which there are several available (Starzl *et al.* 1989b; Rees and Lockwood 1993). Cyclosporine, the most commonly used maintenance drug, is part of a novel group including FK506 and rapamycin, which acts by disrupting signal transduction from T and B cell receptors. Its dosage is limited by its nephrotoxicity, and standard immunosuppressive regimes achieve a lowered cyclosporine dosage by supplementation with drugs which have complementary effects on the immune system. These include the thiopurines azathioprine and 6-mercaptopurine, which may act by inhibiting the primary antibody response; and steroids, which have a widely disseminated effect, especially on cytokines. Since basic homeostatic mechanisms in host defence are compromised by immunosuppressive agents, the risk of opportunistic infections is a major concern in the post-OLT setting. In addition, studies of immunosuppressed individuals showed a 100-fold greater incidence of malignant disease than in the general population. There is therefore an ongoing need to tailor treatment to disease activity.

In Cambridge, a triple regimen of azathioprine, corticosteroids and cyclosporine is administered post-OLT and donors, who are scarce, are not selected by HLA typing (Caine 1993). HLA typing of both donors and recipients is however carried out post-transplantation, for retrospective analysis. Also, only organs from HCV-negative donors are used. In addition to the improved quality of life of liver transplant recipients, their short-term (five-year) survival rates are high, and the prognosis improves markedly after 5 years survival, many recipients being still alive and enjoying good health after 19 years of follow-up (Starzl *et al.* 1989b).

There is very little need for OLT in cases of acute hepatitis, especially when caused by HAV or HEV, where the native liver usually recovers given time. With the more complicated sequelae of hepatotrophic virus infection, however, liver transplantation plays a significant management role. HBV-related fulminant hepatic failure (FHF) is characterised by massive hepatocyte necrosis accompanied by an intense immune response, which leads rapidly to low viral loads. This explains the much lower rate (15%) of HBV recurrence post-OLT for FHF than for chronic liver disease (70%) (Tibbs and Williams 1995). HCV can be transmitted by donor organs, such as heart, liver, kidney, and pancreas, to transplant recipients (Lau *et al.* 1993; Maple *et al.* 1994; Pereira *et al.* 1995). OLT for the treatment of HCV-related end-stage liver disease almost always leads to recurrent HCV infection in the patient (Wright *et al.* 1992; Pons 1995). HCV RNA is detectable in 87% of graft recipients by one year post-OLT (Konig *et al.* 1992), with viraemia often detectable within days of grafting (Maple *et al.* 1994). The course of post-OLT hepatitis C is different for acquired and recurrent infections. Post-OLT acquired HCV infection has a similar profile to PTH in immunocompetent patients, except that antibody response is greatly attenuated in the context of post-OLT immunosuppression, with seroconversion rates of about 40% for newly-infected patients vs. 85-90% in PTH cases (Poterucha *et al.* 1992). This effect is apparently HCV-specific, as transplanted patients remain capable of generating high-titre antibodies against such pathogens as herpes simplex virus (HSV), cytomegalovirus (CMV) and rotavirus (Hsu *et al.* 1994; Maple *et al.* 1994). This is similar to observations in congenitally immunocompromised

patients with PTH, in whom the appearance of anti-HCV antibodies was delayed for over three years (Cornu *et al.* 1994). Recurrent HCV infection post-OLT is indistinguishable virologically and histologically from acquired infection in the same context (Chazouilleres *et al.* 1994; Gretch *et al.* 1995). In both cases, OLT is followed by a dramatic increase in HCV viral titres, up to 16-fold the pre-OLT levels. The absence of allograft damage in some immunosuppressed patients imples that high levels of circulating virions may be tolerated without direct hepatic damage.

The clinical consequences of post-OLT recurrent HCV infection are highly variable, ranging from asymptomatic to a disease tempo exceeding that in *de novo* infection. Three general patterns of recurrent disease occur in this setting:-

- In the short term, recurrent post-OLT HCV infection is usually benign (Shiffman *et al.* 1994; Chemello *et al.* 1995). Unlike recurrent hepatitis B or D, which are usually accompanied by impaired allograft function and serious morbidity, the risk of acute organ damage due to HCV infection is low in most transplant recipients (Martin *et al.* 1991; Konig *et al.* 1992; Pons 1995). The mean time to the first sign of disease is 95 days, and there is evidence of persistently normal graft function in reinfected patients followed up for up to 10 years(Boker *et al.* 1995). The combined evidence indicates that in most patients, anti-HCV seropositivity pre-OLT does not significantly affect graft outcome.
- Approximately 5% of OLT recipients appear to clear HCV, so that sensitive and repeated testing by PCR for the presence of the virus in serum gives negative results (Konig *et al.* 1992; Sallie *et al.* 1994; Boker *et al.* 1995)
- In a third group of patients, recurrent hepatitis C shows a remarkably rapid progression to cirrhosis, with patients requiring re-transplantation within two years (Martin *et al.* 1991; Belli *et al.* 1993; Lerut *et al.* 1995).

Patients with end-stage cirrhosis show significantly lower HCV titres than those with less advanced disease, probably due to the marked decreased in hepatic mass associated with liver disease (Duvoux *et al.* 1995). HCV reinfection, on the other hand, is characterised by high-titre viraemia in the initial post-OLT period, often accompanied by graft damage

(Gretch *et al.* 1995). The significance of this highly efficient replication is unclear. In one study, no correlation was found between RNA titre and histology, implying that post-OLT, HCV-mediated liver damage is not dependent on viral titres alone (Herrero *et al.* 1995). Another report indicated a positive correlation between HCV RNA levels and severity of post-OLT hepatitis C (Zein*et al.* 1995). The possibility exists that the varying results are type-dependent, i.e., different HCV genotypes may have differential pathological profiles. The development of primary liver cancer, for example, has been shown to occur earlier in patients infected with type 1 than type 2 HCV (Boker *et al.* 1995). This may indicate that certain subtypes replicate more efficiently.

The outcome is benign for most OLT recipients. It is nevertheless important to clarify the relationship between types and disease profiles, in order to administer appropriate prophylaxis. The risk of recurrent disease is apparently independent of levels of immunosuppressants administered (Farges *et al.* 1995). Studies have linked the male sex, advanced age and infection with HCV subtype 1b as independent risk factors for hepatocellular carcinoma development (Bruno *et al.* 1995; Marin *et al.* 1995).

Two studies have been carried out on HCV genomic variation following OLT. One study demonstrated the near-identity of paired pre- and post-OLT samples from individual patients, and constituted the first direct evidence of HCV reinfection of previously infected OLT recipients (Feray *et al.* 1992). In the second study, a post-OLT decrease in quasispecies complexity was described for variants of the NCR and E2/NS2 regions (Martell *et al.* 1994). The coincidence of decreased HCV genomic heterogeneity with post-OLT immunosuppression indicated that immune pressure might be the dominant force controlling the dynamic behaviour of HCV in infected individuals. However, the probable involvement of other factors was suggested by the finding, post-OLT, of variants that had not been detected pre-transplantation.

The current study was aimed at clarifying the role of the immune response in immunosuppressed liver transplant recipients. Despite the mild nature of most hepatitis C

cases, the persistent nature of HCV infection constitutes a major clinical problem, exacerbated by the dearth of efficient experimental models or culture systems for studying the virus. Orthotopic liver transplantation, routinely used to treat HCV-related end-stage liver disease, is accompanied by the administration of immunosuppressive drugs. By eliminating or attenuating the host immune response, the post-OLT immunosuppressive regimen offers a model for the study of direct pathogenic effects of the virus on the grafted liver, HCV recurrence being an almost inevitable sequel to transplantation. The fact that transplant recipients are monitored closely means that abundant clinical and laboratory data, from which further evaluations can be made, are generated. These include concentrations of serum AST and ALT, bilirubin, and albumin, derangement of which occurs in liver damage. Hence, the OLT model offers the possibility of clarifying which viral and host factors are important in the pathology of HCV-related disease.

1.6 OBJECTIVES

The aims of this investigation were as follows:

1. To investigate the role of the immune system in the genomic diversity of HCV by comparing mutations in the HVR1 of two groups of infected patients:-
   (a) patients receiving immunosuppressive treatment after OLT;
   (b) patients with chronic HCV-related hepatitis not undergoing any treatment.

2. To attempt to differentiate the respective roles of the host immune response and immunosuppressive treatment in HCV genomic variability.

3. To derive from these studies a better understanding of the mechanisms of host-virus interaction underlying HCV graft infection and HCV genomic variability.

CHAPTER TWO

2.1 MATERIALS

2.1.1 Patient samples

Plasma samples for the preliminary part of this study came from archives held at the East Anglian Blood Centre. The anticoagulants heparin, EDTA or acid citrate dextrose were used to prevent blood coagulation. In addition, whole blood was freshly-drawn from candidates for orthotopic liver transplantation (OLT) at the Addenbrooke's Hospital Department of Surgery, or from asymptomatic HCV carriers referred to the consultant hepatologist at the Department of Medicine. Table 2.1 lists all individuals from whom blood samples were drawn for this study. They fell into three categories: HCV-infected candidates who had received, or were scheduled to undergo liver transplantation following HCV-related end-stage chronic liver disease; asymptomatic blood donors found to be HCV-positive by antibody and PCR tests; and negative controls.

Table 2.1: Patients involved in study

Patient	Age (yrs)	Origin	Sex	Category
A1	59	Italy	M	T x
A2	62	Italy	M	T x
B1	NK	Italy	F	T x
C1	56	Italy	M	T x
C2	NK	Italy	F	T x
C3	58	Italy	M	T x
De	35	Italy	M	T x
F1	NK	Italy	F	T x
L1	66	Italy	M	T x
L2	48	Italy	M	T x
L3	46	Italy	F	T x
M1	52	Italy	F	T x
M2	NK	Egypt	M	T x
W1	NK	UK	F	T x
V	32	UK	M	PC

W	38	UK	M	PC
X	35	UK	M	PC
Y	37	UK	M	PC
Z	24	UK	M	PC
Ad	21	UK	M	Ref
Dix	44	UK	M	Ref
Fra	37	UK	F	Ref
Har	42	UK	M	Ref
Wad	32	UK	M	Ref
D1	NK	UK	NK	NC

*Tx = OLT candidates or recipients; PC = HCV PCR positive controls from blood bank archives; Ref = HCV-positive blood donors from whom nucleotide sequences corresponding to multiple time points were obtained for comparison with OLT-influenced samples. NK = not known, NC = negative control.

2.1.2 Buffer recipes

Compositions for the buffers described in the text are given below.

Buffer	**Composition**
Eco RI buffer (10 x, New England Biolabs))	50 mM NaCl; 100 mM Tris-HCl (pH 7.5); 10 mM $MgCl_2$; 0.025% Triton X-100.
"Basic EDTA"	2 M NaOH; 2 mM EDTA.
Guanidinium buffer	4 M guanidinium thiocyanate; 0.8% ß-mercaptoethanol; 25 mM sodium citrate (pH 7); 1% sarkosyl and 0.2 M sodium acetate (pH 4).
Ligation buffer (10 x, Invitrogen)	60 mM Tris-HCl, pH 7.5; 60 mM $MgCl_2$; 50 mM NaCl; 1 mg/ml BSA; 70 mM β-mercaptoethanol; 1 mM ATP; 20 mM DTT; 10 mM spermidine.
PBS	125 mM NaCl; 25 mM NaH_2PO_4 (pH 7).
PCR buffer (10 x)	50 mM KCl; 100 mM Tris-HCl (pH 8.3); 15 mM $MgCl_2$.

Pst I buffer (10 x, New England Biolabs)	100 mM NaCl; 50 mM Tris-HCl (pH 7.9); 10 mM $MgCl_2$; 1 mM DTT.
RBC lysis buffer	155 mM NH_4Cl; 10 mM $KHCO_3$; 0.1 mM $Na_2EDTA.H_2O$.
SOC bacterial growth medium	1% tryptone; 0.5% yeast extract; 10 mM NaCl; 2.5 mM KCl; 10 mM $MgCl_2$; 10 mM $MgSO_4$; 20 mM glucose; pH 7.0.
SSC stock solution (20x)	3 M NaCl; 0.3 M Sodium citrate
"Salty PEG"	20% PEG; 2.5 M NaCl.
5 x TBE	0.225 M Tris-borate; 0.005 M EDTA.
TE buffer	10 mM Tris-HCl pH 8; 1.0 mM EDTA.
TErase	50 µg/ml RNase A in TE buffer.
Tri-dye	0.25% (w/v) bromophenol blue; 0.25% (w/v) xylene cyanol; 0.5% (w/v) orange G; 50% glycerol; 1 x TBE buffer.
X-gal stock solution	100 mg/ml X-gal in dimethyl formamide.

2.1.3 Bacterial culture media

The basic bacterial growth media 2 x TY and SOC (Sambrook *et al.* 1989) as well as all agar culture plates were obtained from the LMB core facility. Agar plates contained 100µg/ml ampicillin (AMP) and, for NovaBlue cells (Invitrogen) only, 15 µg/ml tetracycline (TET). 2 x liquid medium, containing 100 µg/ml AMP, was used for growing selected colonies.

2.1.4 PCR and sequencing primers

In the design of PCR primers for the NCR and E2 regions, consensus sequences spanning the desired segments were derived from 88 and 45 HCV sequences, respectively, obtained from the sequence database. Primer sequences are given in Table 2.3.

22 HCV ANTIBODY ASSAYS: ELISA & RIBA

Antibody to HCV in serum or plasma samples was detected in routine screening assays carried out at the East Anglian Blood Centre using the ABBOTT HCV EIA 2nd Generation kit (EIA-2), which detects antibodies expressed by putative structural and non-structural regions of the HCV genome. Diluted human serum or plasma samples were incubated with recombinant (*E. coli*, yeast) HCV antigen. Any specific antibody present was affixed to the bead-bound antigen. Unbound material was washed off, and human immunoglobulins detected by incubating the bead-antigen-antibody complex with a solution containing horseradish peroxidase-labelled goat antibodies directed against human immunoglobulins. The substrate for the subsequent enzymatic reaction was o-phenylenediamine.2HCl (OPD). A yellow-orange colour developed in proportion to the amount of bound anti-HCV antibody.

Confirmatory assays for samples positive by EIA-2 were carried out with the Chiron RIBA HCV 3.0 Strip Immunoblot Assay (SIA) kit. This assay was designed as an additional, more specific test on specimens found to be repeatedly reactive using the ELISA kit. Antigens included were recombinant c33c and NS5; and synthetic 5-1-1, c100 and c22, immobilised as individual bands onto test strips. Since recombinant c33c and NS5 were produced as fusion proteins with human superoxide dismutase (hSOD), recombinant hSOD was included as a control band on each strip, to allow detection of antibodies against hSOD which are not specific for the HCV-encoded portions of the recombinant HCV antigens. Low- and high-concentration immunoglobulin internal control bands were also blotted onto each strip. In the standard protocol, the diluted specimen was incubated with the strip. HCV-specific antibodies, if present, bound to the corresponding antigen and/or peptide bands on the strip. Unbound material was removed by aspiration and washing. The strip was then incubated in the presence of peroxidase-labelled anti-human IgG conjugate, which bound to the human IgG portion of the antigen-antibody complex. The colorimetric enzyme detection system involved incubation with hydrogen peroxide and 4-chlro-1-naphthol, to give an insoluble blue-black reaction product at each specific HCV antigen, peptide or control band, giving a visual band pattern on each individual strip.

Anti-HCV reactivity and specificity ws determined by comparing the intensity of each antigen band to that of the human IgG internal control bands on each strip, and visually graded from -1 to 4+. A sample was considered negative if no bands, or if the hSOD band only, was present with a grading of 1+ or higher; indeterminate if a single band, with or without the hSOD band, was 1+ or higher; and positive if two or more hands had a reactivity greater than 1+.

2.3 LEUKOCYTE EXTRACTIONS

Centrifugation is often used for the fractionation of biological material. In isopycnic centrifiguation, particles are separated on the basis of their density. Metrizoic acid (3-acetamido-5-(N-methylacetamido)-tri-iodobenzoic acid, molecular weight 628) belongs to a family of compounds that have been used extensively for density gradient separations. The behaviour of metrizoate solutions in a centrifugal field is intermediate between those of sucrose and CsCl. Thus, sodium metrizoate, unlike sucrose, does sediment to give a density gradient, but at a much slower rate than CsCl. Lymphoprep™ consists of a 9.6% (w/v) Na metrizoate solution containing 5.6% polysucrose, an erythrocyte aggregation polysaccharide. The density of the solution is 1.077 g/cm^3. In the standard leukocyte fractionation procedure, heparin- or EDTA-treated blood, diluted with an equal amount of PBS, is layered over approximately half its volume of Lymphoprep™ to give a discontinuous gradient. After spinning at low speed for about 30 mins, erythrocytes are pelleted, together with granulocytes. Platelets and mononuclear cells, including lymphocytes, monocytes and some basophils, form a distinct band at the interface between the sample layer and the Lymphoprep solution, from where they can be harvested.

2.3.1 PBMC extraction from whole blood samples

Whole blood (10 - 20 ml) was spun down (1,000 rpm 20 min in a Centra-3C centrifuge). The cell fraction, from which the plasma had been carefully drawn off, was made up to the original whole blood volume in PBS. The diluted cells were mixed, avoiding air bubbles, with a 10 ml pipette, and layered gently onto an equal volume of the Ficoll®/sodium metrizoate solution Lymphoprep™ (Nycomed Pharma AS, Oslo, Norway). Leukocytes were separated by spinning at 1,500 rpm for 30 min (rapid acceleration, brake off). The buffy layer, containing leukocytes, was transferred to a

fresh 50 ml tube which was topped with PBS for a first wash. This involved mixing 10 times with a 10 ml pipette, followed by a 1200 rpm, 10 min spin. Two additional washes were carried out in 30 ml and 10 ml PBS. The final wash was carried out in 1 ml PBS. The supernatant was aspirated off, and the cell pellet stored at -40°C.

2.3.2 Isolation of white blood cell subpopulations

A 20 ml volume of whole blood was collected for each patient, in 20 U/ml heparin anticoagulant (CP Pharmaceuticals) and 0.6% (w/v) dextran (M.W. 156,000, Sigma). Following incubation at 37°C for 45 min, the WBC-enriched top layer was eased onto an equal volume of Lymphoprep™, then spun, with rapid acceleration, at 1500 rpm for 30 min. The top layer was aspirated off, and the buffy layer, containing lymphocytes, collected in a fresh 50 ml tube. The third layer was decanted, and the pellet (containing polymorphs) was dissolved in 5 ml RBC lysis buffer. It was then incubated for 5 min on ice. The solutions containing lymphocytes and polymorphs were each made up to 50 ml with PBS, and spun at 1200 rpm for 10 min. Additional washes were carried out in 30 ml, 25 ml and 2 ml PBS. A final spin (2000 rpm in 1.5 ml PBS in an Eppendorf tube) resulted in cell pellets containing predominantly lymphocytes or polymorphs , verified by visualizing under a light microscope, and by an electronic cell counter (Serono System 9000* Diff model). Lymphocyte preparations were 60 - 92% pure (mean 75%), with 6 - 30% granulocyte contamination, while polymorph preparations were 48 - 81% pure, with 7 - 27% (mean 13%) lymphocyte contamination (see appendix C). The cell pellets were stored at -40°C until analysis.

2.4 RNA EXTRACTION

In nucleic acid extraction from prokaryotic and eukaryotic sources, the two major problems are deproteination and inactivation of nucleases. Phenol-, proteinase K- or guanidinium-based protocols are routinely used. The inactivation of nucleases is especially important for RNA, which is rapidly degraded by ubiquitous RNAses, present in sources as varied as glassware, hair and dust (Cox 1968). In the classic protocol, phenol extraction was followed by alcohol precipitation (MacDonald *et al.* 1987). This method was more efficient than the common method involving SDS for cell disruption followed by proteinase K digestion.

The efficacy of guanidinium salts in RNA extraction was first demonstrated in the 1950s. Advantages included the selective denaturation of protein in nucleoprotein complexes - while proteins readily dissolve in the highly electrostatic guanidinium chloride, becoming denatured at 4 M, nucleic acid secondary structure remains relatively intact, with DNA retaining its double helical form even at 70°C. Thus, the method allows both deproteination and RNase destruction in a single step. The high efficacy of guanidinium salt action enables nucleic acid extraction from even nuclease-rich sources such as lymphocytes. The most common and consistently successful methods for isolating pure, intact total RNA are modifications of the original guanidinium thiocyanate method of Chirgwin *et al.* (1979). One such modification involves the coextraction with phenol at decreased pH to remove protein and DNA (Chomczynski and Sacchi 1987). It is often the method of choice when multiple RNA extractions are performed.

2.4.1 Proteinase K-based RNA extraction from plasma samples

To 250 µl thawed plasma sample was added an equal volume of 2 x PK lysis buffer (0.1 M Tris-HCl pH 8.4; 0.02 M EDTA; 0.4 M NaCl; 4% SDS; 2 mg/ml proteinase K). The mixture was incubated at 60°C for 1 h. 500 µl phenol:chloroform (1:1) were then added, and the tubes were revolved for 20 min on a rotating wheel at room temperature. Following a 5-min spin to separate the phases, the aqueous phase was re-extracted once with phenol:chloroform, and once with chloroform alone. RNA in the aqueous phase was precipitated overnight at -20°C in 2 volumes ethanol and 20 µg glycogen carrier (Boehringer Mannheim). It was washed in 70% ethanol, freeze-dried and re-dissolved in 10 µl RNase-free water.

2.4.2 GTC-based RNA extraction from plasma or WBC samples

RNA extraction solution was made up by mixing 250 µl plasma with 500 µl each of water-saturated phenol (Appligene) and guanidinium buffer (4 M guanidinium thiocyanate (GTC), 0.8% ß-mercaptoethanol, 25 mM sodium citrate (pH 7), 1% N-lauroyl sarcosine (Sarkosyl®, BDH) and 0.2 M sodium acetate (pH 4.0). RNA in the aqueous phase was propanol-precipitated overnight at -20°C with 20 µg glycogen (Boehringer Mannheim, Germany). After a wash with 250 µl 80% ethanol and being left to dry at room temperature, the RNA pellet was dissolved in 20 µl DEPC-treated water

(Ultraspec™ water, Biotecx Labs Inc., USA, supplied by AMS Biotech). The solution was incubated 5 min at 70°C to ensure complete RNA dissolution.

25 REVERSE TRANSCRIPTION OF HCV RNA

Traditional RNA detection methods - such as Northern blots and RNA dot/slot blots - require very large amounts of total RNA. RT-PCR provides a more sensitive method requiring much smaller amounts of RNA, and is the only feasible method for HCV, which is present in the circulation of infected individuals only in very low amounts. The cDNA template for PCR is generated by reverse transcription. The reverse transcriptases of both avian myoblastosis virus (AMV) and Moloney murine leukaemia virus (MMLV) have been shown to give comparable results (CLONTECH Laboratories Inc. 1991; Hu *et al.* 1993). To obtain maximal yields, it was important to include human placental RNase inhibitor in the cDNA synthesis reaction.

Priming of HCV RNA for cDNA synthesis in this study was carried out by two methods. In the random priming method, the entire RNA population was converted into cDNA by priming with random hexamers. Two gene-specific PCR primers were then added for PCR amplification. Alternatively, the HCV-specific 3' (antisense) primer was annealed to genomic RNA for site-specific synthesis of the sense strand. Its extension with reverse transcriptase generated a cDNA template for the 5' (sense) primer during subsequent PCR.

2.5.1 cDNA synthesis from RNA obtained by PK lysis

10 µl RNA solution (see 2.2.1) were heated 5 min at 65°C. To the heat-denatured RNA, which was rapidly chilled on ice, were added 10 µl 2 x RT buffer I (0.2 M Tris-HCl pH 8.4 (at 42°C); 0.28 M KCl; 0.02 M magnesium II chloride; 0.8 M dNTPs; 0.02 M DTT; 1.5 µg random primers (Promega) and approximately 14 U each of RNasin® and AMV RT. RNA/primer annealing was carried out by a 10 min room temp incubation; cDNA synthesis at 42°C for 90 min; and RT inactivation at 80°C for

10 min. cDNA samples were stored at -20°C after mixing with 2.5 µg yeast tRNA (Boehringer Mannheim).

2.5.2 cDNA synthesis from RNA obtained by GTC-based extraction

3 µl RNA solution obtained as in 2.4.2 were incubated 30 - 45 min at 42°C in a total volume of 20 µl containing 1 x RT buffer II (0.15 mM dNTPs; 15% DMSO; 1.5 µM antisense oligo; 14 U RNAse inhibitor (RNasin®, Promega); and 10 U MMLV reverse transcriptase (HT Biotech). The stock 10 x RT buffer II used comprised 0.5 M Tris-HCl pH 8.0; 0.5 M KCl; 50 mM $MgCl_2$, 50 mM DTT; and 0.5 mg/ml RNase-free BSA in aqueous solution.

2.6 PCR AMPLIFICATION OF HCV cDNA FROM THE *NCR* AND THE *E1/E2* CODING REGIONS

Thermal cycling parameters were optimized for each primer set and PCR target as follows: Program I of Table 2.3, with its relatively low number of cycles, was found to be sufficient to amplify the HCV NCR, a well-conserved and easily accessible target. Programs V - VII were used in (unsuccessful) early attempts to amplify *E1/E2* sequences based on low-stringency annealing parameters. Program IV was used in the one-step amplification protocol described in section 2.6.3(a) (J. Saldanha, personal communication). Programs II and III, devised by Dr. Juraj Petrik for nested PCR, were found to amplify *E1/E2* more consistently than program I. The relatively high annealing temperature of 55°C (program II) was used in order to minimise non-specific binding, and was lowered slightly for nesting PCR (program III) to permit rapid amplification of specific sequences derived from first PCR. Rigid control over contamination of PCR reactions by HCV sequences was exercised by adherence to the Kwok-Higuchi guidelines where possible (Kwok and Higuchi 1989). Thus, disposable gloves, changed frequently, were used; deionized water used in RNA extraction and reverse transcription reactions was treated with diethylpyrocarbonate (DEPC), and that used for PCR amplification was filter-sterilized or autoclaved; pre- and post-PCR solutions were handled with separate, dedicated pipettors; aerosol-free (plugged) pipette tips were used; reagents were divided

into aliquots to minimize the number of repeated samplings; "premixtures" were formed from reagents before dividing into aliquots, minimizing the number of sample transfers; PCR and cloning procedures were carried out in separate laboratories; DNA was added last to reaction tubes; and non-DNA negative controls were added to each set of reactions. In addition, for the semi-automated protocol described in Section 2.7, where up to 60 samples were processed at one time, dUTP was substituted for dTTP in the first amplification experiment. In subsequent experiments, reaction mixtures were treated with uracil N-glycosylase, which cleaves the dUTP-containing nucleotides carried over from the previous experiment, thereby preventing their use as PCR templates (Longo *et al.* 1990). For this set of experiments, the different stages of the protocol (reagent preparation, RNA extraction, first and second PCR) were carried out in separate, custom-designed hoods.

In nested PCR, a second set of internal primers is used to amplify a portion of the first amplification product. The specificity achieved by this method is equivalent to that from a hybridization reaction - non-specific priming is eliminated by accepting only samples with bands of the precisely correct size on the EtBr-stained gel.

The rapidity and simplicity of the PCR assay was increased by combining the reverse transcription step with the first PCR round in the same reaction mix. This also decreased labour intensiveness and potential for contamination.

Hot-start PCR was employed as an alternative to nested PCR, in order to decrease the risk of contamination due to the transfer from first- to second-round amplification. The reaction consists of an upper and lower portion separated by a wax barrier. The final components only mix when the temperature reaches 80°C. The method had previously been demonstrated to show equivalent sensitivity to conventional nested PCR (Saldanha and Minor 1994).

All PCR, unless otherwise stated, was carried out on a TRIO-Thermoblock (Biometra) or on a Perkin Elmer Cetus DNA thermal cycler.

Table 2.2: *Oligonucleotide primers used for sequencing and nested amplification of HCV genome*

No.	Position[1]	Sense	Sequences	2nd PCR product size (bp)
			Set 1 (NCR):-	
3762	63-82	outer +	TTC ACG CAG AAA GYG YCT AG	
3764	321-339	outer -	TGC ACG GTC TAC GAG ACC T	
3763	83-102	inner +	CCA TGG CGT TAG TAY GAG TG	
3765	295-319	inner -	CAC TCG CAA GCA CCC TAT C	237
			Set 2 (E1/E2):-	
3768	1284-1303	outer +	GGI CAY CGY ATG GCI TGG GA	
3769	1889-1908	outer -	CIR TCG GTI GTI CCC ACB AC	
3766	1300-1319	inner +	GGG AYA TGA TRA ACT GG	
3770	1851-1870	inner -	CAG TAI ACY GGR CCR CAY AC	571
			Set 3 (E1/E2)[2]:-	
3958	1287-1306	outer +	ATA *AAG CTT*[3] CAC GGC ATG GCA TGG CAT AT	
3959	1867-1886	outer -	CAC *GAA TTC* GGG GCT GGG AGT GAA GCA AT	
3960	1293-1312	inner +	GGT *AAG CTT* ATG GCA TGG GAT ATG ATG AT	
3961	1854-1873	inner -	CTG *GAA TTC* AAG CAA TAT ACC GGA CCA CA	581
			Set 4 (E2)[4]:-	
3771	1436-1453	+	GGG IAA YTG GGC IAA GGT	
3778	1436-1453	+	GTG AGC GTG GGC IAA RGT	
9078	1689-1706	-	ATG TGC CAG CTG CCR TTG	

1 Nucleotide positions given correspond to those of HC-J8 (Okamoto *et al.* 1992)

2 Primer sequences for this set were as previously described in the literature (Farci *et al.* 1992)

3 Italicized sequences represent non-HCV restriction enzyme sites for *Eco* RI (GAATTC) and *Hind* III (AAGCTT)

4 Primers in this set were used for bi-directional sequencing of the HCV HVR.

2.6.1 Nested PCR for HCV NCR and E1/E2

A description of all oligonucleotides used in the study (Protein & Nucleic Acid Chemistry facility, University of Cambridge) is given in Table 2.2. The lyophilised primers, of specified absorbance units,

were dissolved in water to give master solutions of about 5 mM. From these, dilutions were made to obtain working solutions of 25 µM (for PCR) or 5 µM (sequencing). For the first round of the nested PCR procedure, 2 µl cDNA prepared as in 2.5.2 were added to a reaction mix to give final concentrations of 0.5 µM each of sense and antisense outer primers; 31.25 µM dNTPs; 1 x PCR buffer (10 x "Cetus" PCR buffer was made up to contain 0.5 M KCl, 100 mM Tris-HCl pH 8.3; and 15 mM $MgCl_2$); and 2.5 U AmpliTaq® (Perkin-Elmer (USA), supplied by Applied Biosystems) in a total reaction volume of 20 µl.

Table 2.3: *PCR programmes for amplification of various parts of the HCV genome*

Name	Pre-heating	Cycling parameters	# of cycles	Final extension time
Prog I	94°C 1 min 30 s	94°C 25 s (denaturation) 50°C 35 s (annealing) 68°C 2 min (extension)	25	10 min
Prog II	94°C 1 min	95°C 25 s (denaturation) 55°C 35 s (annealing) 72°C 1 min 30 s (extension)	35	
Prog III	94°C 1 min	95°C 25 s; 51°C 30 s; 72°C 1 min 30 s	35	9.5 min
Prog IV	94°C 3 min	94°C 30 s; 50°C 30 s; 72°C 1 min	40	3 min
Prog V	95°C 3 min	95°C 20 s; 48°C 25 s; 72°C 30 s	30	7 min
Prog VI	94°C 1 min	95°C 25 s; 37°C 30 s; 72°C 1 min 30 s	35	9.5 min
Prog VII	94°C 1 min	94°C 30 s; 42°C 30 s; 72°C 1 min	35	6 min

PCR reaction conditions were as follows:- denaturation at 94°C for 25 s; annealing at 50°C for 35 s; and elongation at 68°C for 2 min 30s. The cycle was repeated 25 times, followed by a 10 min final extension step. The second round of PCR (PCR2) was carried out using 2 µl first PCR product as

template, in a total reaction volume of 40 µl containing PCR buffer, 0.5 µM each primer; 37.5 µM dNTPs and 1.0 U Amplitaq. PCR conditions were the same as for PCR1, and corresponded to Prog IV of Table 2.3.

2.6.2 Combined RT-PCR and nesting

To 5 µl RNA solution prepared as described in Section 2.4.2 above were added 45 µl of RT-PCR solution, to give a final concentration, in 50 µl aqueous solution, of 0.2 mM dNTPs, 1 x PCR buffer, and 0.5 µM each outer primer, also containing 2.5 U AmpliTaq® polymerase, 10 U reverse transcriptase ("Super RT," HT Biotech Ltd.) and 14 U RNase inhibitor (RNasin®, Promega). The reaction mix was overlaid with mineral oil and RT-PCR was carried out on a thermocycler (Biometra) programmed to incubate the samples for one hour at 42°C for the initial reverse transcription step, then for 1 min 30 s at 94°C to denature the resulting cDNA and inactivate the reverse transcriptase, followed by 35 cycles of amplification as detailed in Prog II of Table 2.3. RT-PCR1 products were stored at -20°C. PCR2 was carried out on 2 µl of DNA template from the first round of PCR, in a total volume of 40 µl containing PCR buffer; 0.5 µM each inner primer; 37.5 µM dNTPs; and 1.0 U AmpliTaq®. PCR reaction conditions used were as given in Prog III of Table 2.3.

2.6.3 Hotstart PCR

(a) Single-round PCR for NCR amplification: For each reaction, a 25 µl aliquot of PCR bottom mix containing 8 mM dNTPs, 1 x PCR buffer, 8 mM magnesium chloride and 2 µM each of forward and reverse primer was topped with 45 µl melted liquid paraffin or a single wax bead (Speci-microsystems). A 3-min, 80°C incubation served to (re-)melt the wax, giving a solid seal upon re-cooling. Onto this seal were dispensed 65 µl of top mix (1 x PCR buffer and 1.0 U AmpliTaq). Final concentrations of components, after the addition of 10 µl cDNA prepared as in Section 2.5.1, were 0.2 mM dNTPs; 1 x PCR buffer; 2 mM $MgCl_2$ and 0.5 µM primers. PCR reaction conditions were as detailed in Prog IV of Table 2.3.

(b) Hotstart RT-PCR & "nesting" for NCR and E1/E2 fragments: 5 µl RNA were added to 20 µl of a bottom mix containing 1 x PCR buffer, 0.4 mM dNTPs, 1 µM antisense outer primer, 4% DMSO and 0.5 µl each of RNasin® and Super RT. The mix was covered with a wax bead as in Section 2.6.3 (a), then reverse transcription was carried out on a PCR machine set to incubate samples at 42°C for 90 min, followed by a 94°C 1-min incubation (for RT denaturation). The (25 µl) top mix contained 1 x PCR buffer, 1 µM outer sense oligo and 2.5 U AmpliTaq®, to give a total combined concentration of 1 x PCR buffer; 0.5 µM each primer; approximately 0.2 mM dNTPs; and 2% DMSO. PCR conditions were as for Progs I and VI in Table 2.2 for amplification of NCR and E2 fragments, respectively. PCR2 was carried on 2 µl of the initial amplification product using the appropriate inner primers and other PCR components In the same concentrations and volumes as described for PCR2 in Section 2.6.1. Cycling parameters were as detailed in Prog IV of Table 2.3.

2.7 HCV DETECTION USING AVIDIN-BIOTIN TECHNOLOGY

The principle of avidin-biotin technology is the remarkably strong interaction between the tetrameric protein avidin and the water-soluble egg white vitamin biotin (Fig 2.1). The K_a of 10^{15} M^{-1} implies a Gibbs free energy value of about 21 kcal mol^{-1}, a staggeringly large value for the noncovalent interaction of a protein with a small molecule (Richards 1990; Wilchek and Bayer 1990)
The interaction is so strong that even biotin coupled to proteins (through its valeric acid side chain) is available for binding by avidin. Streptavidin is a product with avidin-like activity isolated from various *Streptomyces* species.

Batch processing of serial samples obtained from OLT patients in this study was facilitated by the use of a novel, semi-automated method developed in our laboratory by Dr. Juraj Petrik (Petrik, Allain & Pearson, International Patent Application No. PCT/GB96/01768; Petrik *et al.* 1996). The method utilizes the strong interactive force between streptavidin and biotin to selectively isolate HCV RNA from plasma or serum samples. HCV is virtually unique among RNA viruses in having a poly-U (as

opposed to the more common poly-A) stretch of bases at the 3' end of the genome. Hence, a stretch of nucleotides comprising 40 adenine residues and biotinylated at the 5' end (designated dA40) was used to capture positive-stranded, HCV RNA genome. Any background would be eliminated by subsequent amplification using HCV-specific PCR primers.

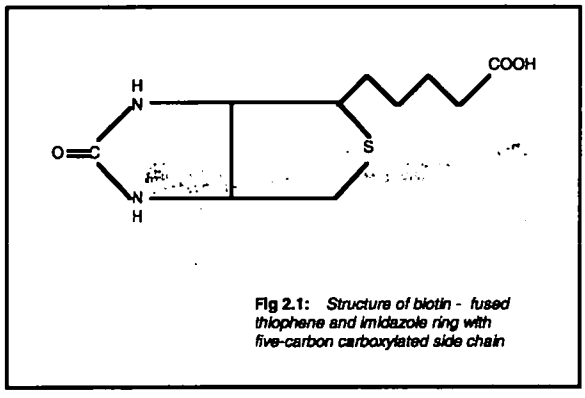


Fig 2.1: *Structure of biotin - fused thiophene and imidazole ring with five-carbon carboxylated side chain*

Southern hybridization of PCR amplicons to nylon membrane (Boehringer Mannheim) was followed by a sinsitive chemoluminescent detection assay involving the use of Lumi-Phos™ 530. This formulation contains Lumigen™ PPD (a stable 1,2-dioxetane) and an enhancer for chemiluminescent detection of alkaline phosphatase. Enzymatic dephosphorylation of PPD produces an unstable intermediate which, as it decomposes, emits blue light (at 477 nm) in direct proportion to the amount of alkaline phosphatase present. The signal is enhanced by the presence of fluorescent micelles formed by cetyltrimethylammonium bromide and 5-N-tetradecanoyl-amino-fluorescein. The fluorescein acceptor emits a bright yellow luminescence at 525 nm, which can be recorded on X-ray film.

All batch extractions and subsequent manipulations were carried out by Gavin Pearson.

Bead preparation: To prepare a stock solution of the capture oligonucleotide, dA40, 5 ml streptavidin-coated paramagnetic particles, designated PMP/S (Promega), were concentrated on a magnetic stand. The supernatant was discarded, and the beads washed twice with 2 ml 0.5 x SSC. After each wash, the supernatant was discarded after magnetic stand concentration of the beads. 5'-biotinylated dA40 (50 pmoles/well) was mixed with 1 ml 0.5 x SSC, and the resulting solution used to

resuspend the PMP/S. The remaining steps were a 3-min room temp incubation, followed by a brief agitation by vortexing; a 0.5 x SSC wash; one wash with 5 x binding buffer (0.05 M Tris-HCl, pH 7.4; 2.5 M LiCl; 0.01 M EDTA); and resuspension in 2 ml binding buffer. The mix was stored at +4°C until needed.

RNA extraction: For each sample, 30 µl of bead-bound capture oligo prepared as described above were mixed with 15 µl 10 x lysis buffer (0.1 M tris-Cl, pH 7.4; 1.4 M NaCl; 0.05 M KCl; 10% triton X-100; and freshly added DTT (50 mM) and RNasin® (20 U)). The 45 µl mixture was added to 105 µl plasma in a single well of an Abbott 60-well diagnostic plate, which was shaken for 10 min on an Abbott Commander dynamic incubator set to 37°C. The particles, with bound HCV, were concentrated using MAG60 (a novel magnetic stand designed for the Abbott diagnostic plate). The supernatant was discarded and the particles were resuspended in 0.5 x SSC and shaken at room temperature. They were concentrated on MAG60, and the supernatant discarded. The wash was repeated, and the sample resuspended in 13 µl DEPC-treated water containing 12.5% DMSO, and incubated for 5 min at 45°C with shaking. The particles were concentrated on MAG60, and 12 µl of eluate were transferred to a System 9600 MicroAmp tube (Applied Biosystems).

RT-PCR: All the RNA obtained as detailed above was used directly in a combined RT-PCR reaction by adding to it 8 µl of reaction mix, to give a 20 µl solution containing final concentrations of 1 x PCR buffer; 0.2 mM dNTPs; 2.25 mM $MgCl_2$; 5 mM DTT; 1 µM each primer; 20 U RNasin®; 8.2 U Super RT; and 1.0 U AmpliTaq®. Reactions were carried out on a GeneAmp PCR System 9600 cycler (Perkin Elmer), programmed to pre-incubate the samples at 37°C for 30 min (for cDNA synthesis) followed by 95°C for 3 min (to destroy RT). Thermal cycling parameters were as for Prog V of Table 2.2. 1 µl aliquots of the PCR1 reaction products were transferred to a second set of 9600 tubes, and 19 µl PCR2 mix (containing PCR buffer, $MgCl_2$, dNTPs, inner primers and AmpliTaq® in the same concentrations as for PCR1) added to each. A 10-min room temp and 10-min 98°C incubation were followed by the same cycling parameters as for PCR1.

Product detection: PCR products were run on a Hybaid Electro-4 Gel system and visualized by illuminating the EtBr-stained gels with UV light. Southern blotting served both to confirm the gel

results and to increase sensitivity where no band was visible by EtBr staining. The Genius non-radioactive nucleic acid labeling and detection system (Boehringer Mannheim) was used. Briefly, the DNA in gel was denatured for 30 min in a 0.5 M NaOH/0.15 M NaCl solution; neutralized by a 30-min incubation in 1.5 M Tris-HCl pH 7.5/0.15 M NaCl; blotted onto Hybond paper for 1 h at about one-tenth atmospheric pressure (75 - 80 mmHg using a Stratagene pressure control station); air-dried for 30 min on Whatman 3 mm paper; wrapped in Saran Wrap and transilluminated on a UV light box for 3.5 min. At this stage it could be stored indefinitely at -20°C. The membrane containing transferred DNA was pre-hybridized in 25 ml buffer comprising 5 x SSC,1% blocking reagent (Boehringer Mannheim), 87.5µl 30% N-lauroyl sarcosine and 50 µl 10% SDS. Hybridization followed in the same buffer containing 10 pmole/ml primer 3792, at 50°C for 5 min. The membrane was washed twice at room temp In 2 x SSC/0.1% SDS, for 5 mln each tlme. Two further 5-min washes at 50°C followed, in 0.5 x SSC/0.1% SDS.

The sequences of the primers used for hybridization are given below:-

3792: **DI**-GGA GAG CCA TAG TGG TCT GC;
4068B: **DI**-CAC TCG CAA GCA CCC TAT C;

where **DI** = digoxigenin.

The digoxigenin-labelled probes were prepared from random hexanucleotides including digoxigenin-conjugated d-UTP by polymerase (Klenow)-mediated transcription using linear denatured DNA of the opposite sense as template strand. Detection was based on the use of Lumiphos-530. For each assay, the membrane carrying the hybridized probe was blocked for 3 h at room temp in 2% (w/v) blocking reagent. It was then gently shaken for 30 min (room temp) with 150 mU/ml high-affinity sheep anti-digoxigenin F_{ab} fragments conjugated to alkaline phosphatase. The membrane was washed twice for 15 min each in Buffer A (0.1 M Tris-HCl, pH 7.5; 1.5 M NaCl) and equilibrated for 2 min in buffer C (0.1 M Tris-HCl pH 9.5; 0.1 M NaCl; 0.05 M $MgCl_2$). The membrane carrying the hybridized probe and bound antibody conjugate was reacted with the Lumi-Phos™ 530 substrate for 30 min at 37°C. It was then exposed to X-ray film (Kodak XAR) to record the chemiluminescent signal.

2.8 VISUALIZATION & EXTRACTION OF PCR AMPLICONS

The fluorescent dye EtBr contains a planar group that intercalates between the stacked bases of DNA. Dye bound to DNA displays an increased fluorescent yield compared to unbound dye (Sambrook *et al.* 1989), which means that small amounts of DNA may be detected even in the presence of unbound EtBr.

2.8.1 Analytical gel electrophoresis of DNA

This was carried out with general purpose agarose having a low coefficient of electro-endoosmosis (BRL ultrapure DNA grade, (-mr = 0.1 - 0.15). Gels were cast at 1.5 - 2%. 50 ml 1 x TBE buffer containing 0.8 - 1 g agarose were heated in a microwave to dissolve the agarose. After the addition of 1 µl stock EtBr solution (at 10 mg/ml), the gel was cast in a gel mould with gel combs in place. Gels were run submerged (Uniscience Minigel System) in 1 x TBE buffer. Samples (usually 10 µl, or 25% of the nested PCR product) were loaded in 1 x tri-dye buffer (5 x stock solution contained 0.25% each of bromophenol blue and xylene cyanol, 0.5% orange G, 50% glycerol and 1 x TBE) and were run at a constant voltage (116V, Atto Crosspower 500) for 17 min, or until the orange G marker dye reached the end of the gel. It was then photographed on a UV transilluminator onto video copy processor paper (Mitsubishi thermal paper) using the GDS 5000 gel documentation system (model P68B, Ultra-Violet Products Ltd).

2.8.2 Preparative gel electrophoresis

This was carried out as for analytical gels, except that LMP (low melting point) agarose (SeaPlaque, FMC, supplied by Flowgen) was used, and dissolution was carried out in 1 x TAE buffer. The gel was run at 75 V for 30 - 35 min in the same buffer, and photographed as described above. Appropriate DNA fragments were excised and purified, where necessary, by agarose digestion and EtOH precipitation.

2.9 RECOVERY AND PURIFICATION OF PCR FRAGMENTS

Purification involved a modification of the method described in the GELase™ (Epicentre Technologies, USA, supplier Cambio) protocol. Gel slices containing the DNA bands of interest were weighed in tared 1.5 ml Eppendorf tubes, and an equal volume UHP water added (1 mg gel being equivalent to 1 µl molten agarose). One-fiftieth volume 50 x GELase™ buffer was then added, to give a total final concentration of 40 mM bis-Tris, pH 6; 40 mM sodium chloride). The gel slice was melted at 70°C for 45 min, and equilibrated 15 min at 45°C. 1.0 U GELase™ enzyme was added, and incubation continued at the same temperature for 1 h or more. 1 volume 5 M ammonium acetate was added, followed by 2 volumes absolute ethanol. DNA was precipitated, in the presence of 20 µg glycogen carrier, by centrifugation at 14,000 rpm for 40 min. It was resuspended in 5 µl water, of which 1 µl was run on an analytical gel to assess yield.

Alternatively, aliquots of the digested agarose solution containing DNA were used directly in ligation reactions.

In cases where only a faint PCR band was visible on the analytical gel, all of the remaining PCR product was first concentrated using a Microcon concentrator (Microcon 30, Amicon) before being run in a preparative gel.

2.10 CLONING OF HCV GENOMIC FRAGMENTS

Many commercial kits are available for the direct cloning of PCR amplicons. All aim at a rapid, one-step strategy for the direct insertion of a PCR product into a plasmid vector. They operate on the fact that Taq polymerase has a non template-dependent activity which adds a single deoxyadenosine (A) residue to the 3' ends of PCR products. Advantages are the elimination of enzymatic modifications of the PCR products; and that the method does not require the use of PCR primers which contain restriction sites. The three kits tested all used the blue/white screening method, whose

molecular basis is the insertional inactivation of the galactokinase gene. The enzyme product of this gene acts on the chromogenic substrate X-gal to give B-galactoside, a compound which imparts a blue colour phenotype to colonies spread on agar plates containing the indicator X-gal in the presence of the enhancer IPTG. 3 TA cloning kits were assessed for efficacy of cloning of the HCV E1/E2 fragment. Except where stated, all materials were supplied with the respective kits. In all cases, the insert DNA was E2 PCR product that had been previously gel-purified as described in Section 2.9. The DNA was roughly quantified, and its size determined, by measurement against known amounts of 123-bp ladder (Gibco-BRL) run on the same gel.

Plasmid DNA was prepared from cultures of cells containing vectors with cloned inserts, and the region corresponding to the E1/E2 junction was sequenced by the Sanger method. Simple minipreps are often contaminated by DNA polymerase inhibitors and small oligos which serve as random primers, leading to "ghost" bands, strong stops and other artefacts. As high-purity DNA templates are essential for efficient sequencing, a protocol involving additional purification steps was simultaneously assessed for its effect on the quality of the indicating autoradiograph.

2.10.1 Cloning of *E1/E2* region of HCV using commercial kits

(a) Invitrogen cloning method: The ligation reactions were carried out according to manufacturers' instructions. For a typical reaction, a total volume of 10 µl was used, comprising 4 U T4 DNA ligase; 1 x ligation buffer; 20 ng pCR™II vector; and about 50 ng insert DNA in sterile water. Ligation was performed overnight, or for a minimum of four hours. In a typical transformation protocol, first 2 µl of 0.5 M ß-mercaptoethanol and then 1 - 2 µl of ligation reaction were added to a vial containing 50 µl thawed OneShot™ competent *E. coli* cells (remaining ligation mixture was stored at -20°C). The cells were incubated on ice 30 min, heat-shocked by a 30 s incubation in a 42°C water bath, then transferred to ice for 2 min. 450 µl pre-warmed SOC medium were added, and the vial containing the cells was shaken for 1 h at 225 rpm on a rotary shaker in a warm (37°C) room. Meanwhile, LB agar plates containing 50 µg/ml ampicillin were prepared by spreading 25 µl X-gal (50 mg/ml) with a glass spreader. The X-gal was allowed to diffuse into the agar for about 1 h. 150 µl from

each transformation vial were plated, using a sterile spreader, on separate, labelled LB agar plates. The plates were inverted and placed in a 37°C incubator overnight. White colonies were picked for plasmid isolation, restriction analysis and sequencing.

(b) Novagen pT7Blue T-Vector Kit Protocol: Ligation mixtures contained 1 µl ligase buffer (200 mM Tris-HCl pH 7.6, 50 mM $MgCl_2$), 5 mM DTT, 0.5 mM ATP, 50 ng pT7Blue(R) T-Vector (@ 50 ng/µl), 2 - 3 Weiss units T4 DNA ligase (New England Biolabs) and 1 - 2 µl GELase'd DNA (0.2 pmol). The mixture was incubated at 16°C 2 h to overnight. 1 µl ligation reaction was added to 20 µl NovaBlue competent cells, which were then chilled and heat-shocked for 40 s as in Section 2.10.1. This was followed by a 2-min ice incubation, then 80 µl SOC were added. Subsequent steps were as for Section 2.10.1 except that the LB agar plates contained the additional antibiotic tetracycline (15 µg/ml).

(c) Stratagene pCR-Script™(SK(+) cloning kit protocol: This one-hour method utilized the cutting properties of the novel restriction enzyme *Srf I*, which recognises the octapeptide sequence 5'-GCCCGGGC-3'. The reaction mix in contained both *Srf* I and T4 DNA ligase (New England Biolabs), which effect cutting and religation, respectively, of the plasmid vector. Incorporation of the PCR fragment into the vector would eliminate the *Srf I* cutting site in an irreversible reaction. At the same time, the dynamic equilibrium between cut and religated vector would be shifted so as to produce more cut vector for amplicon incorporation, thus increasing ligation efficiency. The 11 µl ligation mix contained 10 ng pCR-Script™ SK(+) vector, reaction buffer, 0.5 mM rATP; 5 U Srf I, 1 µl T4 DNA ligase and 50 ng insert DNA. A 2 µl aliquot of the ligation reaction was used to transform 40 µl of Epicurian coli® XLI-Blue MRF' Kan supercompetent cells, pre-chilled on ice in Falcon 2059 tubes after the addition of ß-ME to a final concentration of 25 mM. The mixture was left on ice for 10 min, and was heat-pulsed (45 s) and quenched on ice as above. Cells (50 - 150 µl) were plated onto LB-amp-X-gal-IPTG plates as described above.

2.10.2 PCR screening for recombinant clones (colony screening)

Prior to growing colonies overnight for plasmid isolation, the presence of the appropriate insert was checked by PCR. Universal PCR primers, based on pUC/M13 sequences, were obtained from the LMB and were as follows:-

Primer 41 (reverse primer): d(CAG GAA ACA GCT ATG AC)

Primer 40 (forward primer): d(GTA AAA CGA CGG CCA GT)

PCR analysis was by one of three methods:-

(a) Method 1: Starter culture templates: "Minicultures" were prepared by inoculating 150 µl aliquots of culture medium (2 x TY + AMP) dispensed in microtitre dish wells with individual white colonies using a cocktail stick. The plate was incubated on a shaker maintained at 37°C. After 4 - 8 h, 1 µl of each culture was diluted in 20 µl water. 1 µl of the dilution was added to 25 µl PCR reaction solution comprising 0.2 mM dNTPs, 0.1 µM each of primers 40 and 41, 1 x Promega PCR buffer and 1 U Promega Taq. PCR was carried out in 96-well plates (Falcon 3911) on an MJ Research PTC-100 cycler.

(b) Method 2: Direct colony PCR: PCR was carried out as in 2.10.2 (a) above except that the inoculating toothpick was dipped first in the PCR mix and then in the culture medium. Plasmid DNA was isolated only from those cultures whose corresponding PCR inoculum indicated the presence of the right insert.

(c) Method 3: Novagen method: Each selected colony was transferred, by toothpick, into 55 µl sterile water in a 1.5 ml Eppendorf tube. After briefly vortexing the tube, a 5 µl sample was transferred into 2 ml 2 x TY + AMP and set aside. The tubes containing remaining 50 µl aliquots were floated on boiling water for 5 min and briefly spun in a microcentrifuge (13,500 rpm, 1 min). This was followed by the transfer of 10 µl supernatant into a microtitre dish well or 0.5 ml Eppendorf tube containing 40 µl PCR mix, to give final PCR reaction components as in section 2.10.2 (a).

2.10.3 Preparation & restriction analysis of plasmid DNA

(a) Simple minipreps: These were carried out according to Sambrook *et al.* (1989). Cells from a 3 ml, 18 h culture of One Shot™ cells were pelleted in a 1.5 ml Eppendorf tube (13,000 rpm, 1 min, room temp spin). The cell pellet was resuspended in 100 µl GTE (50 mM glucose; 25 mM Tris-HCl (pH 8.0); and 10 mM EDTA (pH 8.0)). After 5 min at RT, 200 µl soln II (0.2M NaOH, 1% SDS) were added. Mixing by inversion was followed by a 5 min incubation on ice. 150 µl soln III (a mixture of potassium acetate and acetic acid to give a final concentration of 3 M K^+ and 5 M acetate) were then mixed in as before, and the mixture incubated a further 5 min on ice. Cell debris was pelleted by spinning (12,500 rpm, 15 - 20 min) in a refrigerated centrifuge. The supernatant, decanted into a fresh Eppendorf tube, was precipitated with two volumes absolute EtOH. The DNA pellets were dried at room temp for 5 - 10 min then resuspended in 20 µl UHP water.

(b) Phenol-purified plasmid preparations (Novapreps): The steps followed in this protocol were as for simple minipreps in Section 2.10.3(a) up till the addition of soln III and the spinning down of cell debris. At this stage, 300 - 400 µl phenol:CIAA (1:1) were added to the supernatant (where CIAA refers to a 24:1 mixture of chloroform and isoamyl alcohol). The mixture was well-vortexed, and spun 1 min in a microcentrifuge. DNA in the aqueous phase was then precipitated with absolute EtOH and dried as above. Resuspension was in 30 µl TErase (50 µg/ml RNase A in TE buffer), followed by a 15 - 30 min incubation at 37°C. "Salty PEG" (12 µl) was mixed in by pipetting, and the mixture left on ice for at least 1 h. DNA was spun down (13,500 rpm, 10 min, 4°C), the supernatant carefully removed from the (transparent) pellet. and a wash in 70% EtOH carried out. After airdrying, the pellet was resuspended in 11 µl TE.

2.10.4 Restriction analysis of plasmid DNA preps

Restriction digests were carried out on the minipreps in order to ensure that the inserts had been maintained during culture. Enzymes used were *Eco* RI for the pCR™II vector (Invitrogen) and, additionally, *Pst* I for pT7Blue (Novagen) and pCR-Script (IStratagen). A typical reaction mix

contained, in 11 µl total, 1 x buffer, 2 U enzyme, and 1 µl DNA prep, and incubation was at 37°C for 30 min up to overnight. For double digests, *Pst* I buffer was used and the *Eco* RI enzyme component increased to 4 U. All enzymes and buffers were from New England Biolabs. The size of the insert was determined from comparison with an appropriate standard run alongside digestion products on an analytical gel (see Section 2.8.1) after the addition of 2.5 µl tri-dye loading buffer. For simple minipreps, the tri-dye used contained 100 µg/ml RNase A and the samples were incubated, after its addition, at 37°C for at least 15 min before being run on the gel.

2.11 NUCLEOTIDE SEQUENCING

The enzymatic method of Sanger (1981) generates four separate populations of radiolabelled oligos that begin from a fixed point and terminate randomly at a fixed residue. The length of oligos is determined by the location of a particular base along the length of the template. Every base in template has equal chance of being the terminus. Populations are resolved by electrophoresis. When populations are loaded onto adjacent lanes of a sequencing gel, the order of nucleotides along the DNA can be read directly from an autoradiograph of the gel.

Sequenase Version 2.0 is a genetically engineered form of bacteriophage T7 DNA polymerase with a wide tolerance for nucleotide analogs (e.g. dITP) used to resolve regions of compression. The very high processivity and high polymerization rate of this enzyme were exploited by carrying out the reaction in two stages:- (i) a phase involving low dNTP concentrations and low temperature in order to limit the extent of synthesis and ensure efficient incorporation of a radiolabelled dNTP (resulting in primers extended by only 20 - 30 bps) ; (ii) division of each reaction into four tubes, each containing high concentrations of dNTPs and a single ddNTP. This allowed polymerization to continue until a chain-terminating nt was incorporated into the growing chain. The Sequenase®* Version 2.0 DNA sequencing kit (USB) was used for all sequencing.

Polyacrylamide gel electrophoresis was used to separate DNA fragments resulting from the sequencing reaction. The resolving power of this kind of gel is great enough to separate DNA molecules with 1 nt difference. DNA fragments were visualized by autoradiography, which produces a permanent record of the two-dimensional distribution of radioactive atoms on photographic film. {^{35}S} (incorporated into dATP) was used. ß particles emitted by this isotope are known to penetrate film emulsion to a depth of 0.25 mm, sufficient to allow interaction with silver halide crystals therein when film and source were directly apposed. X-ray film was developed in an automatic X-ray film processor (X-Ograph Compact X2).

2.11.1 RNase treatment and denaturation of miniprep DNA solutions

To 20 µl of miniprep DNA was added 1 µl of a 10 mg/ml solution of RNase A (10 µg). After incubating the mixture at 37°C for 20 min, 8 µl of a basic mix (stock solution contained 1 M NaOH and 0.5 mM EDTA) were added. Alkaline denaturation of the DNA (5 min at room temp) was followed by the addition of 4 µl 2 M sodium acetate and 116 µl absolute EtOH. The solution was incubated on dry ice for 10 min then spun down for 10 min at 4°C. It was washed once in 75% EtOH, and resuspended in 15 µl UHPW.

2.11.2 Chain extension reaction

For a typical sequencing reaction, 5 pmoles of primer were added to a 3 µl aliquot of alkaline denatured, RNase-treated DNA prepared as described above. The mixture was heated in a 65°C water bath for 2 min, and allowed to cool slowly (over 30 - 45 min) to below 30°C. It was pulsed briefly and chilled on ice. Labelling mix (7.5 µM each of dGTP, dCTP and dTTP) was diluted 5-fold to working concentration. To the ice-cold DNA/primer hybrid (10 µl) were then added 1 µl 0.1 M DTT, 2 µl of diluted labelling mix, 0.5 µl of {^{35}S}dATP and 2 µl of a 1/6 dilution, in enzyme dilution buffer (10 mM Tris-HCl, pH 7.5; 5 mM DTT; 0.5 mg/ml BSA) of Sequenase version 2.0. Extension was carried out at 17°C for 2 - 5 min.

2.11.3 Chain termination reaction

To 2.5 µl of each termination mix, dispensed into wells of a microtitre plate (Falcon 3911) or 0.5 ml Eppendorf tubes and pre-warmed to 37 - 42°C, were transferred 3.4 µl of the extension reaction. The mixture was incubated a further 3 - 5 min at the same temperature. Reactions were stopped by adding 4 µl of Stop Solution (0.05% each of bromophenol blue and Xylene Cyanol FF, 20 mM EDTA and 95% formamide), and stored at -20°C until run.

2.11.4 Casting, loading, running and autoradiography of sequencing gels

Sequencing reaction products were subjected to electrophoresis on denaturing polyacrylamide gels containing 6 M urea using vertical gel tanks (Cambridge Electrophoresis Ltd). Gels were cast between 20 cm x 40 cm or 20 cm x 50 cm glass plates separated by 0.4-mm thick spacers (Bio-Rad). The top plate was treated with silicone solution to allow its easy detachment from the gel at the end of the run. A 50 ml gel solution was made up by adding 20 ml UHP water to 23 g urea (Gibco-BRL), followed by 150 µl of 10% ammonium peroxodisulphate (ammonium persulphate - BDH), 5 ml of 10 x TBE and 7.5 ml of 40% acrylamide (Severn Biotch Ltd - stock solution contained 40% acrylamide and 2.105% bis acrylamide). Just before pouring the gel, TEMED (100 µl, Sigma) was added to the solution. The gel was left to dry, with sharkstooth comb in place, for at least an hour. Samples were heated to 75°C for 5 min immediately before loading (1.5 µl of each) onto lanes created by the sharkstooth comb.

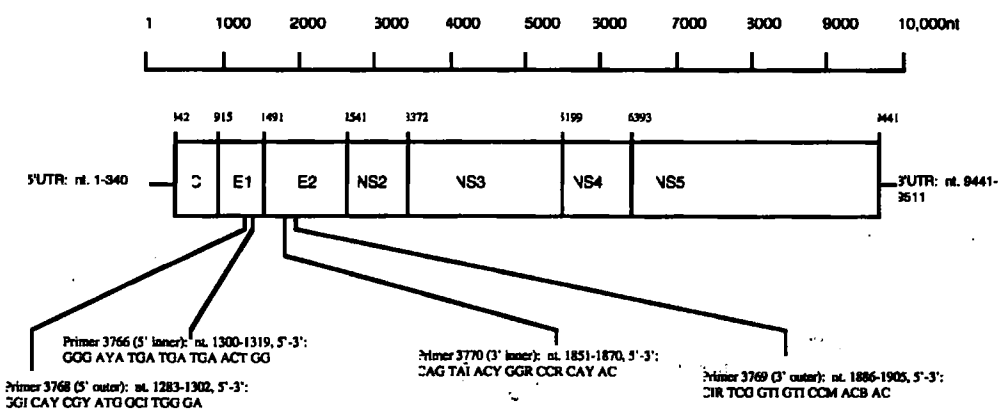
Gels were run at constant power (37 W) for 90 min (40 cm gel) or 140 min (50 cm gel), or until the BPB band reached the end of the gel. The plates were carefully separated and the spacers removed. The plate to which the gel remained attached was balanced, gel side up, on 25 mm-high supports (empty racks from the Sarstedt Tip Stackpack were used) located in a shallow tray. Enough fixing solution (10% each of methanol and acetic acid in water) was dispersed onto the gel to cover it completely (50 - 60 ml). After 5 - 15 min, the liquid was drained off by tilting the plate, and the fixing process repeated with a fresh layer of fixing solution. Excess fluid was removed from the surface of the gel, which was then dried onto a piece of Whatmann 3 MM paper using a commercial

dryer (Atto Rapidry, supplied by Genetic Research Instrumentation). Saran wrap placed round the gel on the paper before drying prevented spread of radioactivity.

2.12 ANALYSIS OF SEQUENCING RESULTS

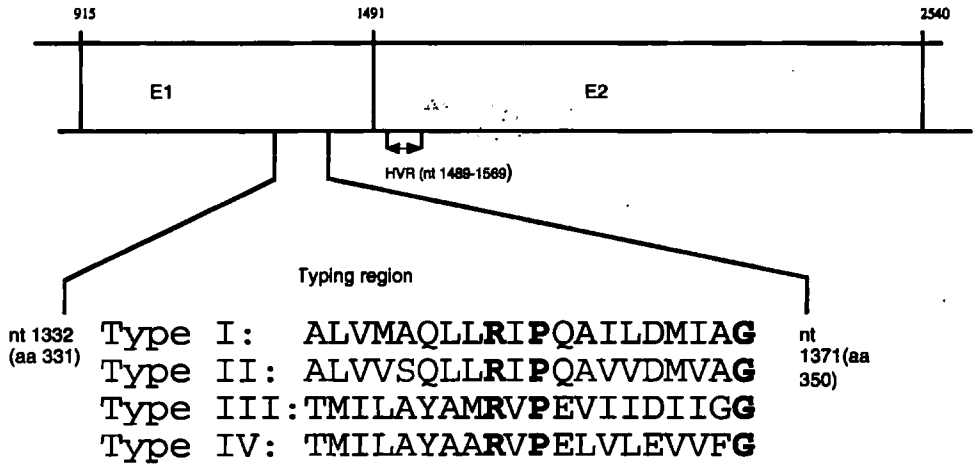
Nucleotide sequences were aligned and checked using the IBI sequence analysis software programs AssemblyLIGN™ and MacVector™ (International Biotechnologies Inc). Typing was based on a previously-described amino acid sequence homology scheme (Okamoto *et al.* 1992; Bukh *et al.* 1993). Figure 2.1 shows primer sequences and locations for the E1/E2 PCR product encompassing HVR1. Type-specific sequence homologies contained within the amplified region enabled a simultaneous determination of the HCV sub-type (figure 2.2). The PCR product was cloned into a bacterial vector, and sequenced by the dideoxy nucleotide chain termination method.

Figure 2.1: *HCV genomic organisation showing location of E1/E2 PCR primers*



Coding regions for putative proteins shown as C (core); E1 and E2 (envelope), NS2-NS5 (nonstructural). UTR = untranslated region. Amplified fragment (371 bp) includes HVR (nt. 1488-1608). I = inosine; Y = T or C; R = A or G; M = A or C; B = not A. Numbering corresponds to that of HC-J8 isolate (Okamoto *et al.*, Virology 188: 331-341 (1992)).

Figure 2.2: *Type-specific amino acid sequence homologies in 3' E1*



Part of HCV genome showing type-specific amino acid homologies within PCR product. One-letter amino acid code depicts residues peculiar to the various subtypes. Absolutely conserved amino acids in boldface. Ref: Okamoto *et al.*, Virology 188: 331-341 (1992).

The nucleotide sequence diversity in each set of sequences was calculated as the mean of the total number of mutated nucleotides between all possible combinations of two sequences. The ratios of replacement to silent mutations (R/S ratios) were calculated for clones corresponding to each sample, and the significance of conservative vs non-conservative amino acid replacements was analysed.

Statistical analysis: Fisher's exact probability test was used in all statistical analysis.

CHAPTER THREE: OPTIMIZATION OF REACTION PROTOCOLS

The research which formed the basis of this thesis involved the generation and amplification of HCV cDNA from a large number of samples, as well as the sequencing of the *E1/E2* region from multiple clones derived from individual samples. Special attention was therefore given to optimizing efficiency and effectiveness at every stage of the procedure. The optimization of reaction protocols led to the discovery of two atypical phenomena, described in this sections 3.2 and 3.3.

3.1 RNA extraction & RT-PCR

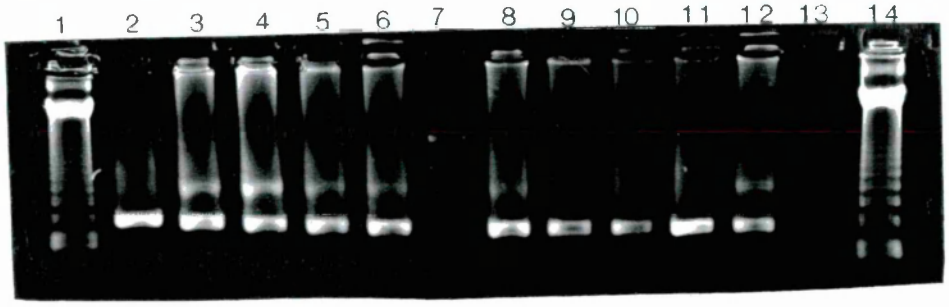
The two methods of HCV RNA extraction used in the study employed either proteinase K or guanidinium thiocyanate as the denaturing agent. The guanidinium-based protocol gave the most consistent results. In amplifying sequences from the HCV NCR and *E1/E2*, RNA was reverse-transcribed to give cDNA, which was then subjected to two rounds of amplification. Alternatively, the reverse transcription reaction was combined with the first PCR round in a single tube, after which second PCR was carried out. The two protocols gave identical results, as shown in Fig 3.1. In addition, for the *E1/E2* region, only the protocol involving a combination of reverse transcription and PCR gave positive results. Hot start PCR is described in chapter 2. In an alternative protocol, assessed because of its relative simplicity, all reaction components were assembled on ice, then transferred rapidly to the PCR block at denaturation temperature, for the amplification reaction. The results of both protocols were identical single, distinct PCR bands with no smearing, as depicted in Fig 3.2. An absence of cross contamination was demonstrated by sequencing of the "fingerprint" HVR1 region.

3.2 Effects of heat or chemical denaturation on amplification of HCV RNA

It was discovered that the difference between a positive and a negative amplification signal, for the *E1/E2* PCR product only, was critically dependent on heating the RNA at 70°C for a

relatively long period before the combined RT-PCR step. Four different heat denaturation regimes were compared to test this effect.

Fig. 3.1: *Results of PCR of the HCV non-coding region - a comparison of separate PCRs and combined RT-PCR1*



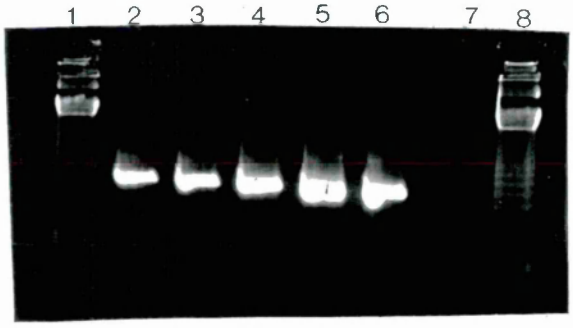
Lanes 1, 14 = 123-bp ladder; lanes 7, 13 = negative controls (plasma from non-HCV-infected donors); lanes 2 - 6 = cDNA derived by separate RT and nested PCRs; lanes 8 - 12 = cDNA derived by combined RT-PCR1 and nested PCR. Note that the cDNAs in lanes 8 - 12 are the same as those in tracks 2 - 6, respectively.

In an experiment involving 12 isolates, RNA was extracted from 250 µl plasma, and dissolved in 20 µl water. 5 µl aliquots from each sample were then subjected to RT-PCR as described in section 2.6.2, except that in three cases (i.e., for each of the remaining 5 µl aliquots) the RNA was first heated at 70°C for 1 h, 65°C for 5 min or 90°C for 5 min. Plasma samples used in this experiment came from asymptomatic blood donors (Ad, Har, Wad); HCV-infected liver transplant recipients (Cap, DK, GS); and individuals who had been repeatedly seropositive for HCV by first and second generation ELISA and RIBA, as well as by PCR of the non-coding region (V, W, X, Y and Z). Gio was a patient seropositive for HCV by both screening and confirmatory tests (see section 2.2), but repeatedly negative by PCR of the non-coding region. D was an isolate from a non-carrier of HCV, used as negative control in all extraction/amplification reactions. Results are shown in Table 3.1(a), and indicated that heating the RNA for 1 h at 70°C (regimen II) was considerably more effective than any of the other treatments.

To further assess the effect of denaturation on the outcome of RT/PCR and nested PCR, RNA was extracted from 6 samples (including 1 negative control, D), in duplicate. The procedure was as given in section 2.4.2, except that for one set of samples, the RNA was

dissolved in 4 μl of the strong denaturing solvent formamide instead of 20 μl water (i.e., in each case, 25% of the RNA obtained from 250 μl plasma or serum).

Fig. 3.2: *Results of PCR of the E1/E2 region by combined RT-PCR1 and nested PCR*



Lanes 1, 8 = 123-bp ladder; lanes 2 - 6 = 571-bp product of *E1/E2* amplification from 5 different plasma samples; lane 7 = negative control

RT-PCR was carried out on 5μl water-dissolved or 1 μl formamide-dissolved RNA, and was preceded by heat-denaturation as in regimen II of Table 2.2 for the aqueous solution. The tube containing the 1 μl RNA solution in formamide was floated for 35 s or 45 s in a boiling water bath, then snap-frozen on ice before RT-PCR, which was carried out in a total volume of 20 μl containing the same reagents in the same concentrations as in Section 2.6.2. Results, shown in Table 3.1(b), indicated that although the 70°C heating step still gave the best outcome, heating in formamide for 35 s was a viable option. The heat-denaturation period could be decreased to 30 min without affecting results, and the majority of the amplification reactions in this study were accordingly carried out with an initial RNA denaturation of 30 min at 70°C.

3.3 Effects of heparin on amplification of HCV sequences from plasma

Mg^{2+} concentration is known to be a critical factor affecting PCR outcome. In PCR mixes containing otherwise optimal amounts of Mg^{2+}, it is often necessary to increase Mg^{2+} concentrations where EDTA is present in the sample, due to the chelating effect of EDTA on divalent cations. In a preliminary study to determine the relative effects of the anticoagulants heparin and EDTA on amplification of HCV RNA, six different plasma samples were subjected, in duplicate, to RNA extraction and RT-PCR as described in

sections 2.4.2 and 2.6.2. The samples had been collected in either heparin or EDTA, and the protocol followed was otherwise identical for all 12 initial plasma aliquots. It was found that bands of the relevant size (the region amplified was a 237-bp segment of the NCR) were obtained only from heparinised samples. In contrast, no PCR bands were obtained when RT/PCR and nested PCR of the *E1/E2* region were carried out on aliquots from the same samples. Experiments had been carried out in exactly the same way for both sets of experiments, except for the specific PCR primers used. Hence, heparin was found to inhibit HCV amplification differentially, depending on the region of the genome targeted. This difficulty was resolved by replacing all heparinised samples with serum samples obtained from similar time points, or by using plasma derived from citrate-preserved blood, for *E1/E2* PCR.

Table 3.1: *Results of amplifying HCV E1/E2 fragments from various isolates using different RNA denaturation regimes. "Patient samples" refers to isolates from asymptomatic carriers or liver transplant recipients; "Controls" = isolates routinely used as positive or negative controls (D = single negative control); "+", "-" = positive, negative E1/E2 PCR result, respectively. Programs I - V refer to different RNA denaturation regimens, where I = no denaturation; II = 70°C 1 h; III = 65°C 5 min; IV = 90°C 5 min; & V = RNA dissolved in formamide before heating in boiling water for 45 s (a) or 35 s (b).*

(a) *Comparison of 4 heat-denaturation programmes*

Pro-	Patient samples						Controls					
gram	Ad	Cap	DK	Gio	GS	Har	V	W	X	Y	D	Z
I	-	-	-	-	-	-	-	-	-	+	-	+
II	-	-	+	-	-	-	+	+	+	+	-	+
III	-	-	-	-	-	-	-	-	-	+	-	+
IV	-	-	-	-	-	-	+	-	-	+	-	-

(b) Comparison of heat- and chemical-denaturation effects on RT-PCR of HCV RNA: PCR of E1/E2 region.

	Ad	Har	Wad	V	W	D
II	+	+	+	+	+	-
V(a)	-	+	+	-	-	-
V(b)	+	+	+	+	-	-

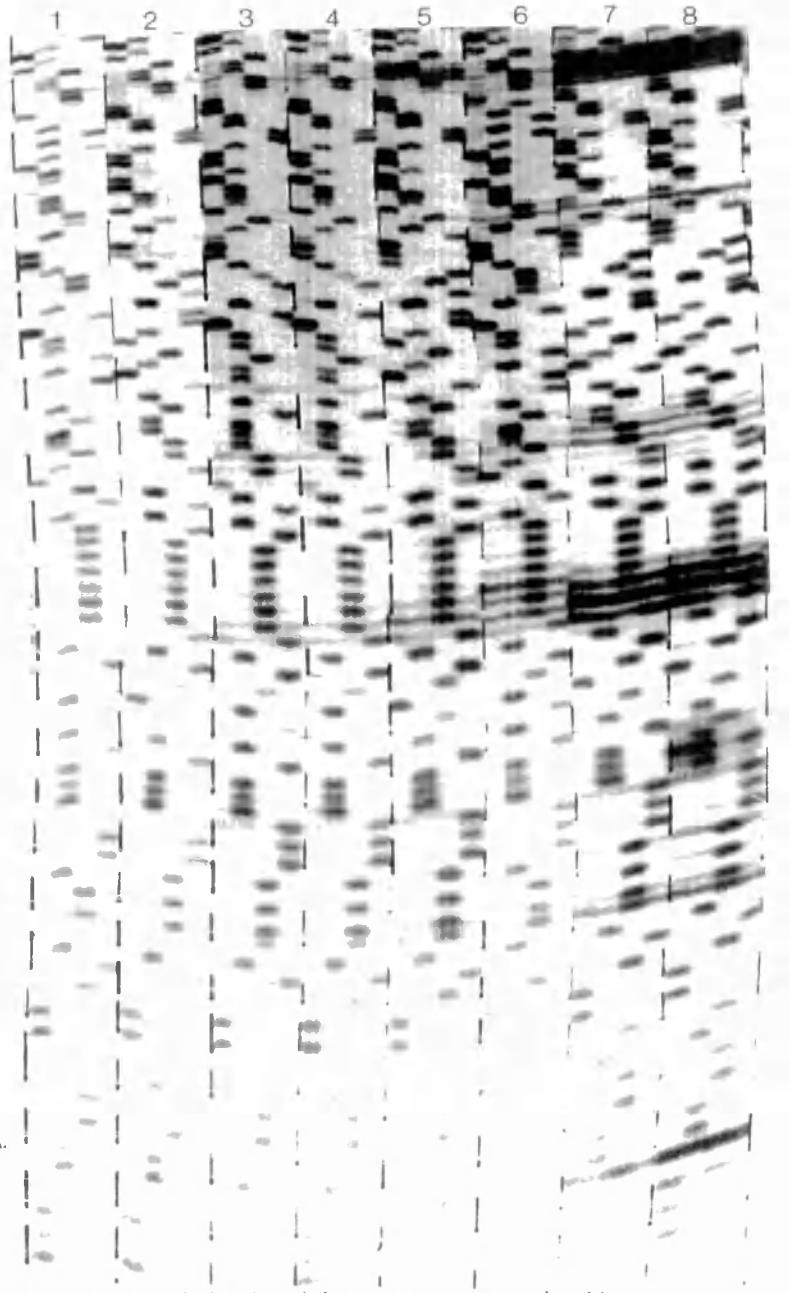
The results depicted in Table 3.2(b) were obtained from fresh plasma aliquots of each sample, different from the aliquots of 3.2(a). Note that samples Ad and Har are PCR-positive with program II on this, as well as on subsequent, occasions.

3.4 Cloning and sequencing of PCR products

Three commercial kits were compared for efficacy of cloning PCR products. The 571-bp product of *E1/E2* PCR was cloned into vectors provided with the Invitrogen, Stratagene and Novagen kits according to manufacturers' instructions (see Section 2.10.1). Only Invitrogen showed consistent incorporation of the PCR fragment, with a false positive level of virtually less than 1% for the white colonies. For this reason, it was possible to dispense with the colony screening assay (Section 2.10.2) for all recombinant clones. Throughput time from PCR to cloning was also decreased by carrying out the cloning reaction with the amplified DNA still in the GELase-digested gel, without further purification.

The Wizard™ PCR preps DNA purification system for rapid purification of cloned plasmid DNA (Promega) was compared with the simple minipreps protocol of Section 2.10.3(a). Similar yields were obtained with both systems, although more uniform bands were obtained with Wizard™ preps. When using the vacuum manifold according to manufacturer's instructions, the Wizard™ preps were two-thirds faster to produce, and purer than those obtained by the simple protocol. Surprisingly however, these preps gave a much poorer read-though when equivalent amounts were sequenced, as depicted in Figure 3.3.

Fig 3.3: *Sanger sequencing of the HCV HVR1: Effect of DNA preparation method on sequencing read-through*



Lanes were loaded in the order:TCGA.
Nos. 1 - 4 = simple minipreps.
Nos. 5 - 8 = Wizard™preps

The relative advantages of Wizard™ preps and simple minipreps are summarised below:-

	Advantages	Disadvantages
Wizard™ preps	- faster - purer preps	-more expensive - more labour-intensive - poorer read-through on sequencing
Simple minipreps	- cheaper - less labour-intensive - better read-though on sequencing	

3.5 Choice of methods

The method used for RNA extraction is an important concern when optimizing HCV PCR. Direct comparison of different isolation methods for HCV RNA showed that of the 4 commercially available kits for nucleic acid isolation (RNAzol B, TRISOLV and ULTRASPEC for RNA and IsoQuick for total nucleic acid), RNAzol B was the most efficacious (Nolte *et al.* 1994). In the author's experience, more consistent results were obtained with guanidinium buffer prepared as in Section 2.4.2 than with RNAzol B.

In PCR, problems often arise from the annealing of primers to each other (primer dimer formation) or to non-target areas of the template under non-stringent conditions, giving rise to amplification of spurious sequences, and is especially problematic in the presence of low amounts of starting template. This happens because Taq DNA polymerase can act outside its optimal temperature of 72°C, retaining significant activity even at room temperature. In hot start PCR, a subset of the reaction components is assembled to lie under a wax barrier. A key constituent is added to the top of the barrier such that all components only mix when the temperature is above the annealing temperature (i.e., when the wax melts). This method has been found to significantly reduce the formation of non-target PCR products (Saldanha and Minor 1994). In the current investigation, it was found that, despite the low viral titres associated with HCV, a hot start technique was not necessary to ensure highly-specific amplification of HCV DNA fragments, as long as all reaction components were assembled on ice, and were transferred rapidly to the pre-heated heating block at the denaturation temperature. Clear bands of the right size with no smearing were consistently obtained by this method, which was simpler, faster and more convenient than hot start.

Heparin has been described as a potent inhibitor of both Taq DNA polymerase and reverse transcriptase (Willems *et al.* 1994). The mechanism of this inhibitory effect is unclear. Heparin may act by directly blocking the enzymes reverse transcriptase and DNA polymerase. Alternatively, it may interact with DNA or RNA to prevent transcription. Divalent cations such as Mg^{2+} may mediate the interaction between heparin and nucleic acid,

both negatively charged macromolecules. In human population genetic studies involving amplification of blood-derived DNA, the time between blood collection and DNA extraction was found to be critical, implying an interaction between heparin and DNA as a mechanism of inhibition (Satsangi *et al.* 1994). Although the effects of heparin could be reversed by treatment of affected DNA or RNA with heparinase, results were often inconsistent (Izraeli *et al.* 1991; Satsangi *et al.* 1994).

Despite the data on the inhibitory effects of heparin, there are reports of successful amplication of HCV NCR sequences from heparinised blood (Takehara *et al.* 1992; Ouanian *et al.* 1995). Indeed, the preliminary experiments described in section 3.2 indicated that, when amplifying the HCV NCR, more consistent results were obtained with heparinised than with EDTA-containing blood. The author's results with amplification of different areas of the HCV genome suggest that the inhibitory effect of heparin may be template-dependent to some extent, in particular that the NCR is, perhaps anomalously, resistant to this effect. The latter view is buttressed by evidence of peculiar secondary structure at 5' end of HCV genome that might contribute to the decreased capability of heparin-binding (Brown *et al.* 1992). Magnesium ion concentration is known to have profound effects on PCR (Innis *et al.* 1990), which may explain the initial negative PCR results obtained for HCV NCR amplification described in section 3.2. It is possible that chelation of Mg^{2+} ions by EDTA altered PCR conditions sufficiently to prevent the generation of visible products from the non coding region.

In summary, these preliminary experiments led to the consistent use of:-

- serum samples, or plasma obtained from non-heparinised (usually citrate-preserved) blood, for PCR of the *E1/E2* region enclosing the HVR1;
- a heat denaturation temperature of 70°C, applied for 30 - 60 min, for denaturing extracted HCV RNA prior to RT/PCR; and
- simple minipreps of section 2.10.3(a) for sequencing using the Sequenase kit.

4.1. PCR of the non-coding region (NCR) and E2 region

Blood samples were obtained prospectively from 14 patients with HCV-associated, end-stage liver disease who were scheduled for orthotopic liver transplantation (OLT) at Addenbrooke's Hospital (see also Table 2.1, category "Tx" patients). All were seropositive for anti-HCV antibody by ELISA and RIBA assays. Following blood separation into plasma and leukocytes as described in chapter 2, section 2.3, HCV RNA extraction was carried out on plasma (or serum) samples by the methods described in section 2.4. cDNA synthesis and nested PCR, or combined RT-PCR followed by a second round of amplification, were carried out as described in section 2.6. In addition, batch processing of serial samples was performed by G. Pearson, using the methods of section 2.7. Results are shown in Table 4.1.

Table 4.1: *Results of PCR of the NCR (N) and E1/E2 region (E) of 5 OLT candidates when heparinised (hep) or non-heparinised (NH) blood was used as HCV nucleic acid source.*

Patient	Pre-OLT		Months post-OLT													
			1		2		3		6		12		18		24	
	N	E	N	E	N	E	N	E	N	E	N	E	N	E	N	E
A1 (hep)	–	–	–	–	–	– [+]	–	–	–	–	–	–	–	–	–	–
C1 (hep)	+	–	+	–			–	– [+]								
(NH)		+						+						+		+
F1 (hep)	+	–	+	+?	+?	–	+	–								
L1 (hep)		–	–		–		+	–								
(NH)		+						+								
L2 (NH)	N.A.			+		+		+		+		+				+
L3 (NH)		+		+												

"+" and "-" signs indicate, respectively, positive and negative PCR results. N.A. = not available. [+] indicates that HCV sequences were amplified from PBMC for that sample.

- 6 of 14 patients (A1, B1, C2, C3, De and M2) were repeatedly negative for serum or plasma HCV by PCR of the non coding region. In A1, however, HCV was isolated from PBMC. It should be noted that this patient, originally seropositive for HCV, had undergone two previous OLTs for HCV-related disease over a three-year period.
- 2 patients did not receive transplants because liver disease was too advanced (A2) or due to the lack of a suitable donor (W1). Although the initial sample was PCR-positive for both the NCR and the *E1/E2* region, no further samples were drawn after the first.
- 1 patient (M1) was PCR-positive pre-transplantation, and at 193 days post-OLT. However, only heparinised samples were available for the pre-OLT time points (see chapter 3, section 3.2). Although the NCR could be amplified from these, attempts to amplify *E1/E2* did not succeed, despite heparinase treatment of the RNA obtained.
- In 1 patient (F1), non-HCV sequences were obtained from PCR-amplified, cloned cDNA corresponding to the *E1/E2* region (see section 4.1.5).

Positive *E1/E2* PCR bands were obtained for serial samples from 3 of the remaining 4 patients (C1, L1 and L3), each set including a pre-OLT sample. The last of these 4 patients (L2) was a candidate for retrospective analysis of HCV genomic evolution in 10 serum samples kindly provided by P. Evans. Although no pre-OLT sample was available for L2, all 10 post-OLT samples, obtained between 31 and 435 days post-OLT, were positive for HCV *E1/E2* sequences by PCR. cDNA bands corresponding to 5 of these time points were cloned, and nucleotide sequences were obtained for the first hypervariable region (HVR1), together with corresponding sequences from C1, in an analysis of the long-term effects of immunosuppression on HVR1 sequence variability, discussed further in chapter 5.

The mean time to reinfection was 67±62 days, although the high standard deviation reflected the variability in this parameter among the 6 patients (C1, F1, L1, L2, L3 and M1) considered in the analysis.

4 patients (C1, L1, L2 and L3) were included in the study. None had any identifiable risk factor for HCV infection. Samples were also obtained from 3 apparently healthy blood

donors (Wad, Fra and Dix) who had a past history of intravenous drug abuse (IVDA), and provided the data for non-transplanted controls.

4.2 Liver function in OLT recipients and asymptomatic subjects as determined from serum ALT values

Serum alanine aminotransferase (ALT) concentrations range from 7 - 50 IU/ml in normal, healthy individuals. For all patients described in Section 4.1.1 above, serum ALT values corresponding to several time points post-OLT were obtained from clinical records at the Addenbrookes Hospital transplantation ward. Data on 8 asymptomatic HCV-infected blood donors were obtained from the East Anglian Blood Centre (EABC), for samples obtained between 1991 and 1993. 7 were confirmed positive for anti-HCV, while the eighth (Dix) was indeterminate. An ALT value above twice the upper limit of normal was considered to constitute liver dysfunction. Figures 1 to 4 depict serum ALT and bilirubin concentrations plotted against number of days post-OLT for C1, L2, L3 and L1.

Patient C1: This was a 56-year-old man who had his first transplant in March 1994 for HCV-related end-stage cirrhosis complicated by hepatocellular carcinoma. Serum ALT rose from 79 IU/ml 11 days post-OLT to 195 IU/ml by one month post-OLT (figure 4.1a). Values peaked at 57 days (319 IU/ml), before falling gradually over the next three months, stabilising at 94 IU/ml by 135 days post-transplantation. Although fluctuating, ALT values remained between 43 and 95 IU/ml for the next year. Deteriorating liver function tests, followed by graft failure, led to a second transplant in patient C1, carried out 386 days after the first.

After the second liver transplant, serum ALT in patient C1 remained much higher and for a longer period than after the first transplant, as shown in figure 4.1b. Serum bilirubin concentrations, also depicted, showed a profile parelling that of ALT levels, except that peaks of bilirubin lagged behind the ALT peaks. Although serum ALT had fallen from 493 IU/ml on the day of transplantation to 144 IU/ml by day 7, it remained quite high throughout the period of follow-up. The 390 IU/ml value recorded 125 days post-OLT2 was

more than three times higher than that taken at a similar time interval from the first transplant (116 IU/ml at 127 days post-OLT1). The patient subsequently died of graft failure due to HCV infection.

Figure 4.1(a): *ALT, bilirubin and HCV PCR profiles for patient C1 after 1st transplant*

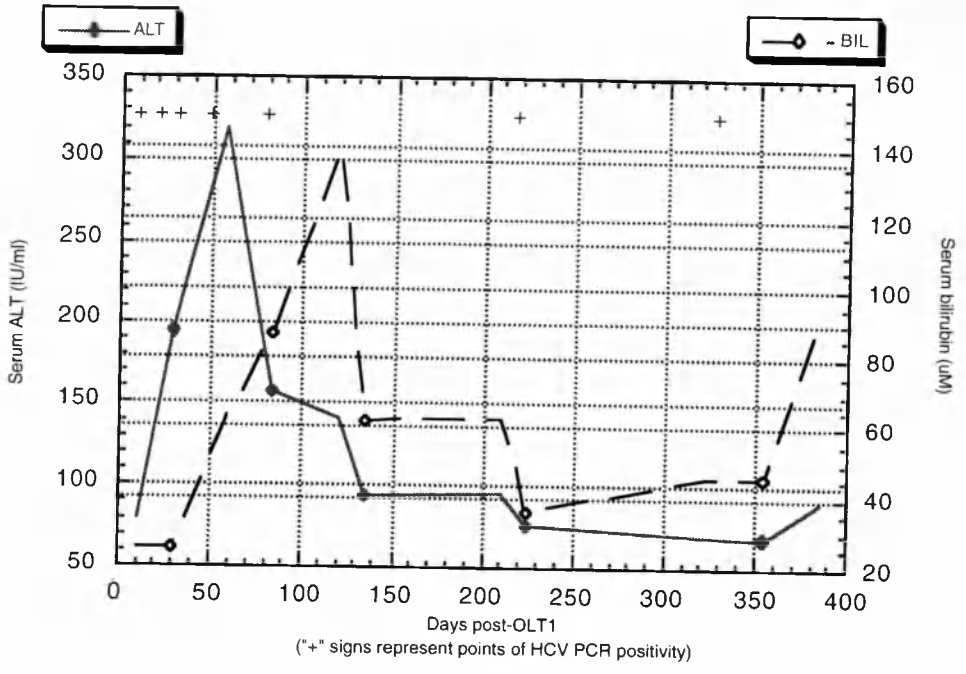
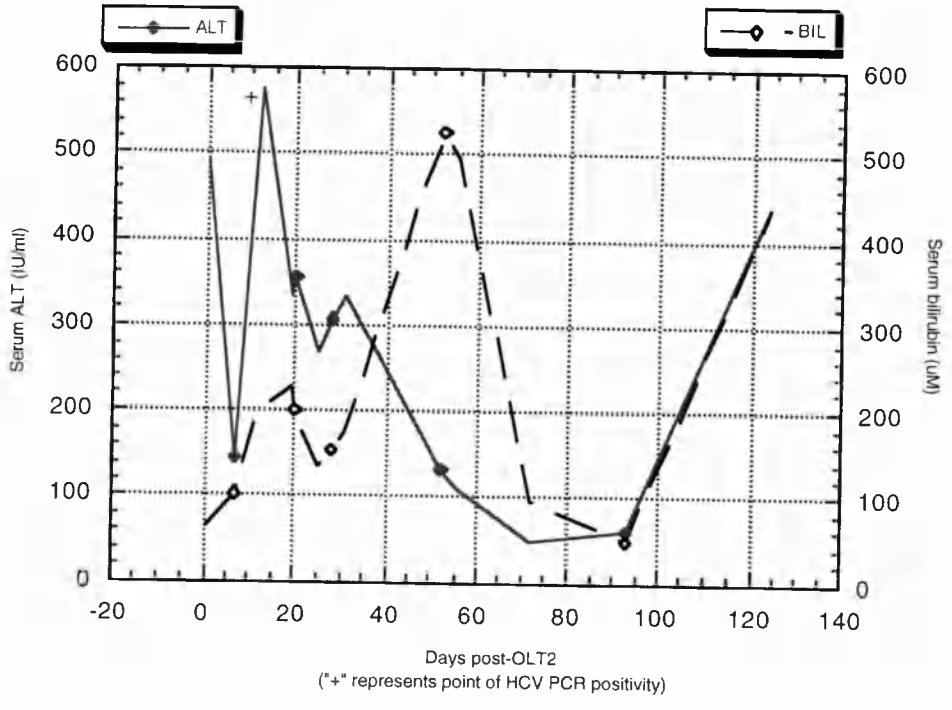
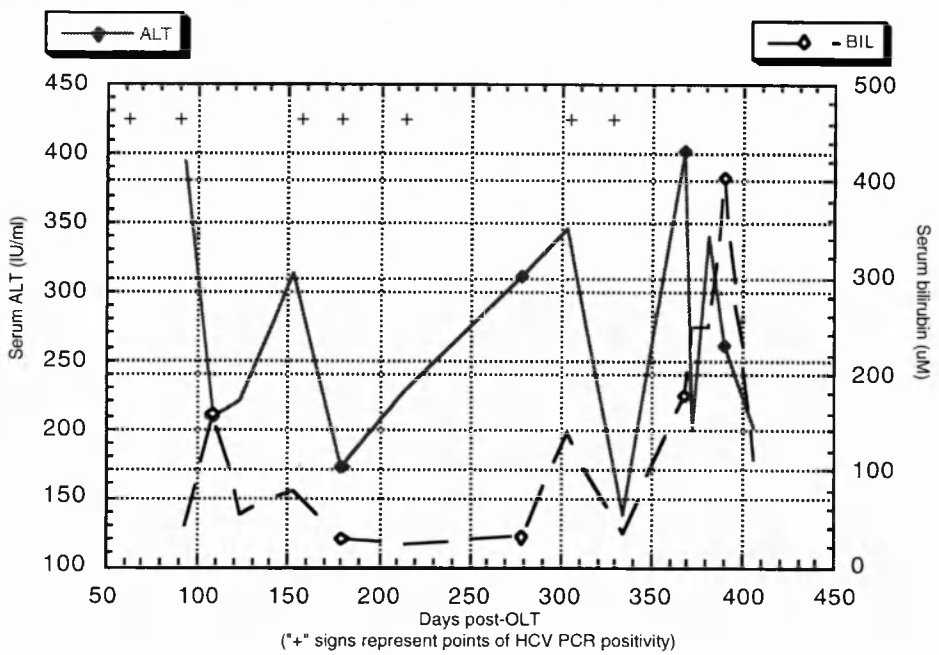


Figure 4.1(b): *ALT, bilirubin and HCV PCR profiles for patient C1 after 2nd transplant*



Patient L2: This was a 48-year-old man who, like patient C1, received two liver grafts. ALT and bilirubin levels in the serum of L2 are shown in figure 4.2. 99 days after the first transplant (OLT1), a high ALT level (395 IU/ml) was observed. ALT levels remained high over the next year, with peaks above 300 IU/ml in the month preceding graft failure due to graft infection with HCV (figure 4.2a). After the second transplant (OLT2), the biochemical course of disease was poor, with ALT peaking at 1300 IU/ml after 40 days (figure 4.2b). Over the next two weeks, serum ALT concentrations fell rapidly, reaching a value of 81 IU/ml within 6 days, and with a mean value of 63±11 IU/ml thereafter. The corresponding values for bilirubin, however, showed a rapid increase. The mean bilirubin concentration of 292±124 μM was an almost 10-fold increase over its mean value (30±12 μM) in the 2 weeks before the ALT peak. The patient died of chronic hepatitis C two months after the second transplant.

Figure 4.2a: *ALT, bilirubin and PCR profiles for patient L2 after 1st transplant*



Patient L3: This 46-year-old woman underwent liver transplantation for HCV-related cirrhosis in February 1994. Data were obtained for the first five months post-OLT, as shown in figure 4.3. A serum ALT value of 153 IU/ml at 7 days post-OLT had normalised by

50 days post-OLT by 14 days post-OLT. Values were fluctuating (92±29 IU/ml) over the next 4 months. At 145 days, when the last record was taken, serum ALT concentration was 65 IU/ml. The patient remained clinically well 12 months post-transplantation.

Figure 4.2b: *ALT, bilirubin and PCR profiles for patient L2 after 2nd transplant*

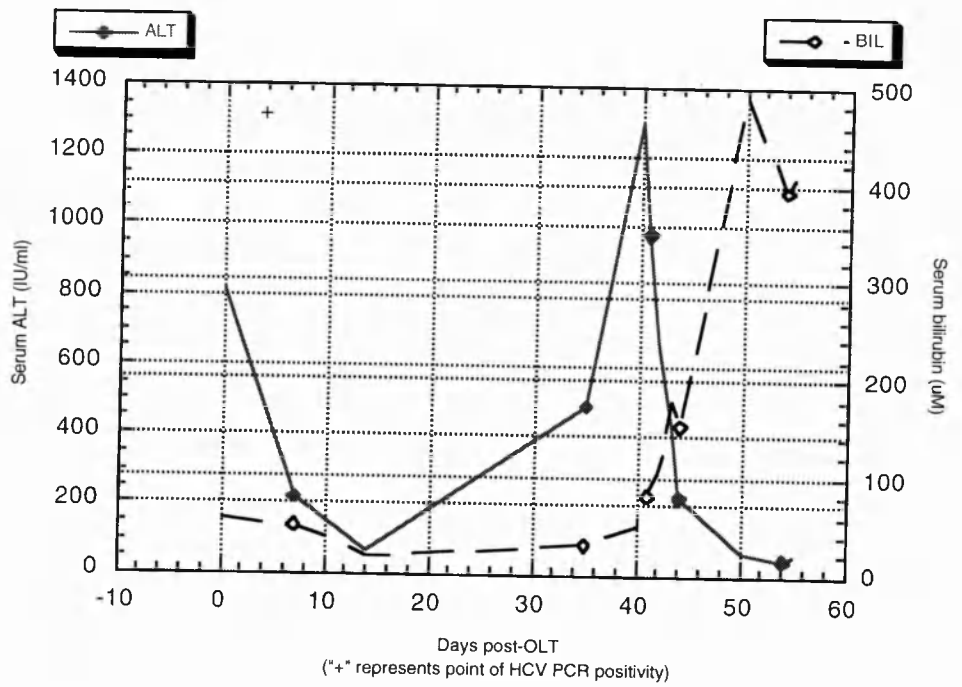
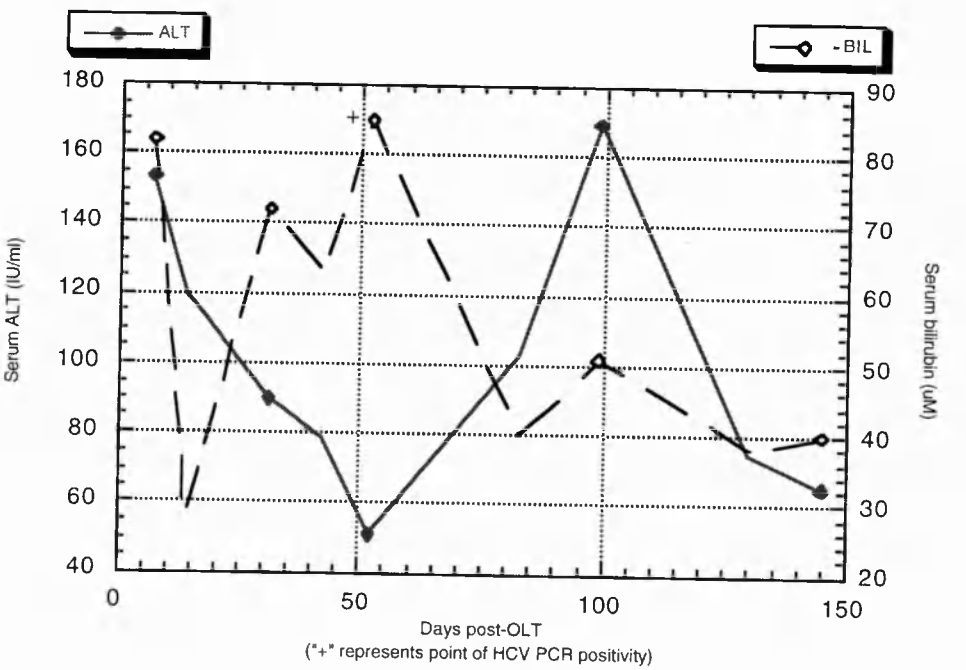
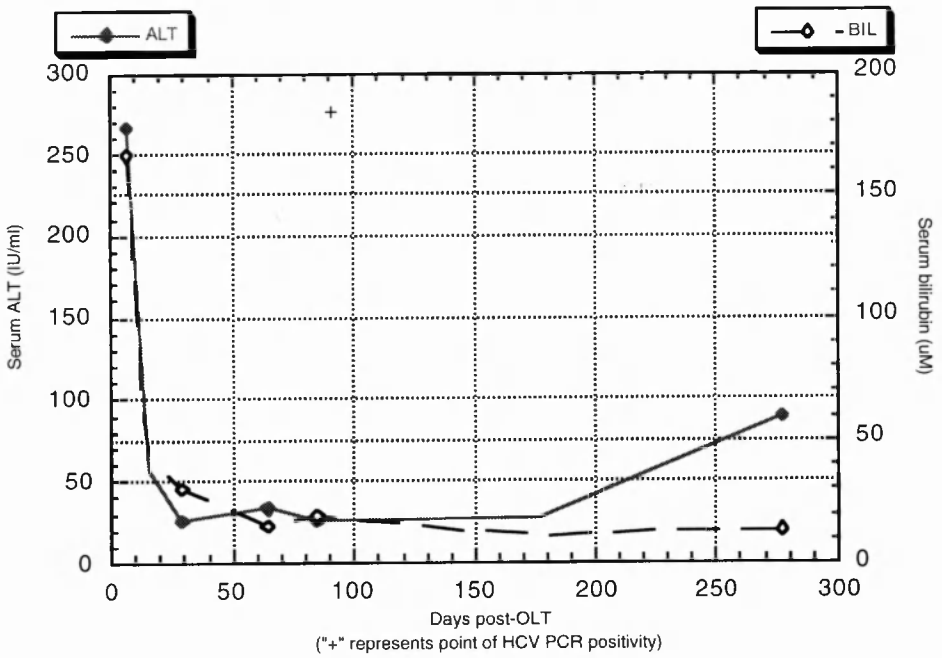


Figure 4.3: *Post-transplantation ALT, bilirubin and PCR profiles for patient L3*



Patient L1: Patient L1, a 66-year-old man, underwent OLT in April 1994 for HCV-related, end-stage cirrhosis. 7 days post-transplantation, ALT concentration remained relatively high in this patient, but had fallen to normal (57 IU/ml) by two weeks post-OLT (see figure 4.4). It remained within the normal range at 1, 2, 3 and 6 months post-OLT. The last serum ALT value recorded (89 IU/ml at 277 days post-OLT) was within acceptable limits. The patient remained well two years post-OLT.

Figure 4.4: *Post-transplantation ALT, bilirubin and PCR profiles for patient L1*



By way of contrast, Table 4.2 shows ALT values for 8 blood donors discovered to be HCV-infected during routine blood screening at the EABC. ALT values were mildly elevated in all, with values from 2 to 4 times the upper normal limit of 50 IU/ml.

Detailed sequence analysis was carried out on three apparently asymptomatic blood donors, described below:-

Patient Wad: This 46-year-old male blood donor had HCV-related chronic hepatitis, detected in February 1993. Although only moderately symptomatic at the biochemical level, liver biopsy revealed chronic active hepatitis with mild fibrosis (stage 2).

Patient Dix: This is a 42-year-old male blood donor found HCV antibody positive in November 1993. He was cryoglobulinaemic, and experienced fatigue and arthralgia.

Patient Fra: This is a 37-year-old woman found HCV antibody positive in a blood donation obtained in September 1994. ALT levels remained persistently normal, but liver biopsy revealed mild, chronic active hepatitis with moderate liver fibrosis (stage 3).

Table 4.2: *ALT levels in asymptomatic, HCV-infected blood donors*

Subject	Age	ALT (IU/ml)
Ad	32	100
Wad	22	≤50
Dix	41	75
Z	22	159
X	33	184
Y	35	109
V	32	194
W	38	141

The mean age of the non-transplanted patients above was 32 (S.D. = 6), compared with a value of 55±7 for the six OLT recipients (A1, C1, L1, L2, L3 and M1) from whom HCV HVR1 sequences were obtained.

4.3 Genotyping of HCV sequences

HCV genotype was determined in 13 infected individuals (6 asymptomatic blood donors and 7 OLT candidates) by the Bukh-Okamoto method, which is based on amino acid sequence homologies exclusive to the distinct types and subtypes (Bukh *et al.* 1993; Okamoto *et al.* 1992). See figure 2.2 and appendix A for a comprehensive list of region-dependent amino acid sequence variations as reported by Okamoto *et al.*

The stretch of 26 - 43 amino acids used for typing was deduced from nucleotides 1332 - 1457 of the prototype HCV isolate HCV-1 (Choo *et al.* 1991), located at the presumptive 3'

terminus of the *E1* gene (see section 1.5.6). It included a 10-residue motif (amino acids 342 - 351) previously shown to be sufficient by itself for HCV genotyping (Bukh *et al.* 1993). In 6 individuals - Ad, Har, Wad, Z1, L3 and W1 - the isolated HCV fragment was classified as Okamoto type I. This corresponds to subtype 1a of Simmonds *et al.* (1995).

Figure 4.5: *Nucleotide (a) and amino acid (b) sequences used for genotyping HCV type I.*

(a) Alignment of reference sequence HCV-1 (nts 1332 - 1454) against homologous sequences from presumptive HCV type I/1a isolates

```
          331                 335                 340                 345
           |                   |                   |                   |
           Ala Leu Val Met ala Gln Leu Leu Arg Ile Pro Gln Ala Ile Leu
      1332                                                          1376
HCV-1*    GCG TTG GTA ATG GCT CAG CTG CTC CGG ATC CCA CAA GCC ATC TTG
Har       ... ... ... G.A ... ... ... ... ... ... ... ... ... ... ...
Wad       ... C.. ... GC. ... ... ... ... a.. G.. ..g ... ... ... ...
Ad        ... ... ... G.A ..c ... ... ... ... ... ... ... ... ... ...
Z         ... C.. ... G.A ... ... ... ... a.. G.. ..g ... ... ... ...
W1        /// //. ... G.A ... ... ... ... a.. G.. ..g ... ... ... ...
L3        /// //. /// /// /// /// /// /// ... ... ..g ... ... ... C..

                    350                 355                 360
                     |                   |                   |
          Asp Met Ile Ala Gly Ala His Trp Gly Val Leu Ala Gly Ile Ala
      1377                                                          1421
HCV-1     GAC ATG ATC GCT GGT GCT CAC TGG GGA GTC CTG GCG GGC ATA GCG
Har       ... ... ... ..c ... ... ... ... ... ... ... ... ... ... ...
Wad       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Ad        ..t ... G.. ... ..c ... ... ... ... ... ..a ... ... .G. ...
Z         ... ... ... ..c ... ... ... ... ... ... ..a ... ... ... ...
W1        ... ... ... ..c ... ... ... ... ... ... ..a ... ... ... ...
L3        ..t ..c ... ..a ..g ... /// /// /// /// /// /// /// /// ///

                    365                 370
                     |                   |
          Tyr Phe Ser Met Val Gly Asn Trp Ala Lys Val Leu
      1422                                                      1457
HCV-1     TAT TTC TCC ATG GTG GGG AAC TGG GCG AAG GTC GCG
Har       ... ... ... ..c ... ... ... ... ... ... ... ...
Wad       ..c ... ... ... ... ..c ... ... ... ... ... ...
Ad        ... ... ... ... ... ... ... ... ... ... ... ...
Z         ... ... ... ... ... ..a ... ... ... ... ... ...
W1        ... ..t ... ... ... ... ... ... ... ..a ... ...
```

(b) Typing of HCV by amino acid sequence comparisons

```
          330       340                 350                 360
           |         |                   |                   |
HCV-1      ALVMAQLLRIP         QAILDMIAGA          HWGVLAGIAYFSMVGN
Har        ...V.......         ..........          ................
Wad        ...A.....V.         ..........          ................
Ad         ...V.......         ......V...          ........V.......
Z          ...V.......         ......V...          ................
W1         //.V.......         ......V...          ................
L3         ////////...         .....I....          ////////////////
```

* Reference sequence = HCV-1 (Choo *et al.*, 1991). For the remaining sequences, only points of nucleotide sequence divergence from HCV-1 are shown. Upper case letters indicate that changes would give a replacement of the encoded amino acid(s).

Figure 4.5 shows the nucleotide (a) and deduced amino acid (b) sequences of the 6 type I/1a isolates, aligned against the corresponding region in HCV-1. By comparison with amino acids 336 - 361 (nucleotides 1330 - 1425) of the type II isolate HCVJK1G (M. Honda, 1991, unpublished; EMBL accession # X61596), HCV derived from 5 isolates (Har, Wad, Ad, Z2 and W1) was classifiable as type II (subtype 1a). The sequence alignments are depicted in figure 4.6.

Figure 4.6: *Nucleotide (a) and amino acid (b) sequences used for genotyping HCV type II.*

(a) Alignment of reference sequence HCV-JK-1G(nts 1330 - 1425) against homologous sequences from presumptive HCV type II/1b isolates

```
          336             340                 345                 350
           |               |                   |                   |
           Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly  Ala
         1330                                                          1377
HCVJK-1G* CAG TTA CTC CGG ATC CCA CAA GCC GTC GTG GAC ATG GTG GTG GGG  GCC
Fra       ... ... ... ... ... ... ... ..t ... ... ... ... .C. ...  ...
L1        ... ... ... ... ... ... ... ..t ... ... ... ... .C. ...  ...
M1        ... ... ... ... ... ... ... ..t ... ... ..t ... .C. ...  ...
C1        ... ..g ... ... ... ... ... ..t A.. ... ..t ... .C. ...  ...
L2        ... ... ... ... ... ... ... ..t ... A.. ... ... .C. ...  ...

                    355               360                 365
                     |                 |                   |
          His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn
        1378                                                          1425
HCV-1     CAC TGG GGA GTC CTG GCG GGC CTC GCC TAC TAT TCC ATG GTG GGG AAC
Fra       ... ... ... ... ... ... ... ..t ... ... ... ... ... ... ... ...
L1        ... ... ... ... ... ... ... ..t ... ... ... ... ... ... ... .G.
M1        ... ... ... ... ... ... ... ..t ... ... ... ... ... ... ... ...
C1        ... ... ... A.. ... ... ... ..t ... ... ... ... ... ... ... ...
L2        ... ... ... ... ... ... ... ..t ... ... ..c ... ... ... ... ...
```

(b) Typing of HCV by amino acid sequence comparisons

```
                340             350             360
                 |               |               |
HCVKJ-1G  QLLRIP        QAVVDMVVGA      HWGVLAGLAYYSMVGN
Fra       ......        ........A..     ................
L1        ......        ........A..     ................
M1        ......        ........A..     ........V.......
C1        ......        ..I.....A..     ................
L2        ......        ...M....A..     ////////////////
```

* Reference sequence = HCVJK-1G (M. Honda, 1991; EMBL accession # X61596). For the remaining sequences, only points of nucleotide sequence divergence from HCVJK-1G are shown. Upper case letters indicate that changes would give a replacement of the encoded amino acid(s).

Criteria for type designations were as follows:-

Type I/1a has the invariant residues QAILD from positions 341-435, respectively. The criteria for designating an isolate as type II/1b included the presence of Val at position 148 **and** at either position 345 or 346. Type IV/2b had a distinct amino acid sequence pattern between residues 342 and 351.

The HCV strain in Dix was genotyped as type IV/2b. Nucleotide and amino acid sequence alignments are shown (fig 4.7) for its coordination with reference type IV sequence HC-J7 (Okamoto *et al.* 1992). Amino acids 331 - 360 (nucleotides 1332 - 1421) were compared.

Figure 4.7: *Nucleotide (a) and amino acid (b) sequences used for genotyping HCV type IV.*

(a) Alignment of reference sequence HC-J7 (nts 1331 - 1421) against homologous sequences from a presumptive HCV type IV/2b isolate

```
          331              335                  340                  345
           |                |                    |                    |
           Thr Met Ile Leu Ala Tyr Ala Ala Arg Val Pro Glu Leu Val Leu  Glu
        1332                                                            1379
HC-J7*     ACC ATG ATC CTC GCC TAT GCC GCT CGT GTT CCT GAG CTA GTC CTT  GAG
Dix        ... ... ... ... ... ..c ..a ... ... ... ... ... ... ... ...  ..a

          347          350                355                  360
           |            |                  |                    |
           Val Val Phe Gly Gly His Trp Gly Val Val Phe Gly Leu Ala
        1380                                                    1421
HC-J7      GTT GTC TTC GGC GGC CAT TGG GGC GTG GTG TTT GGC TTG GCC
Dix        ... C.. ... ... ... ... ... ..t ... ... ... ... ... ...
```

(b) Typing of HCV by amino acid sequence comparisons

```
           330        340                350                360
            |          |                  |                  |
HC-J7      TMILAYAARVP        ELVLEVVFGG        HWGVVFGLA
Dix        ...........        ......L...        ..........
```

* Reference sequence = HCV-J7 (Okamoto *et al.*, 1992). For Dix, only points of nucleotide sequence divergence from HC-J7 are shown. Upper case letters indicate that changes would give a replacement of the encoded amino acid(s).

Analysis of the number and types of changes in the region upon which typing was based revealed a high degree of intratype conservation. The 5 type I/1a isolates differed from the prototype by 4 - 9 (3 - 7%) of the 121 - 126 nucleotides sequenced. A similar degree of divergence was found in the types II/1b and IV/2b isolates. The differences were exclusively

base substitutions, with transitions occurring overall 3.8 times as frequently as transversions. Where amino acid divergence occurred, the replaced residue was usually conserved in character. Amino acid sequence diversity ranged from 2.7 to 3.3%.

In contrast, nucleotide and putative amino acid sequence variability between the reference sequences representing the three (sub)types of figures 4.5 - 4.7, ranged from 27 - 60% and 40 - 80%, respectively. The difference between types was greater than that between subtypes. Thus, the HCV-1 (type I/1a) nucleotide and deduced amino acid sequences differed from those of HCVJK-1G (type II/1b) by 27% and 40%, and from HC-J7 (type IV/2b) by 60% and 80%, respectively.

4.4 Intra-genotype HVR1 sequence diversity

The sequences from the genotyping region contrasted sharply with amino acid sequence divergence in the hypervariable region 1 (HVR1), which lies about 60 residues downstream of the region in *E1* used for typing, from positions 384 - 410.

Figure 4.8: *Deduced amino acid sequences of the dominant HVR1 variant in several HCV isolates. (Absolutely conserved residues are emboldened. Underlined Gly residues were invariant in most (85%) of the sequences).*

```
                        10          20
                        |           |
HCV-1[1]ETHVTGGSAGHTVSGFVSLLAPGAKQN

Har: Q.YVS..SQARGASIITSLFSP.AQ.N
Wad: T.YTS..SAARDTAGLASLFNL.PK.T
Ad:  E.HVT..VAAHATRGLSSLFSL.AQ.N
L3:  G.YTT..AAGQTTSGLAGLFNS.AR.H

1G[2]: G.YVSV.HASQTTRRVASFFSP.SA.K
Fra: Q.HTT..AVARNTYGNTALLTR.PS.Q
L1:  Q.HTV..VAGSTTSGFASLFKL.PS.K
M1:  G.YVT..ASGRTVHGFTGLFSL.PL.T
C1:  T.YVT..SAARRTQHVTSIFSF.SS.N
L2:  Q.YVT..ASARNTHGIASLFAF.PA.K

J7[3]: S.QVT..QAAHTVRGVASIFSP.SR.D
D1:  T.YTT.AQMGRGIFGFSNLFNL.SQ.K
```

1 HCV-1 (Choo *et al.*, 1991)
2 HCVJK-1G (M. Honda, 1991; EMBL accession # X61596).
3 HC-J7 (Okamoto *et al.*, 1992).

Figure 4.8 shows the HVR1 amino acid sequences for the types I/1a, II/1b and IV/2b isolates discussed above. A consensus HVR1 sequence was derived by determining the amino acid occurring with the highest frequency at each of the 27 HVR1 amino acid positions. The consensus sequence is shown in table 4.3, as well as the % frequency of occurrence in the 13 samples analysed:-

Table 4.3: *Consensus amino acid sequence for HVR1 derived from 13 HCV isolates*

HVR1 position	1	2	3	4	5	6	7	8	9	10	11	12
Consensus	G	T	Y	V	T	G	G	S	A	A/G	R	T
% frequency	31	100	62	69	69	92	92	31	62	46	54	38

HVR1 position	13	14	15	16	17	18	19	20	21	22	23	24
Consensus	T	S	G	F	A	S	L	F	S	L	G	P
% frequency	54	31	70	38	46	77	77	92	46	46	100	38

HVR1 position	25	26	27
Consensus	S/Q	Q	N
% frequency	23	100	33

3 of the 27 HVR1 residues (Thr-2, Gly-23 and Gln-26, where the figures refer to the relative positions within HVR1) were consistently conserved in all 13 isolates compared. The amino acid Gly was usually (11 of 13 isolates) found in positions 6 and 7 of the HVR1. However, amino acid sequence divergence between any two isolates was never less than 41% in this region, and reached 67% in several intratype comparisons. There was no significant difference between intra- and inter-type variation in this region. The degree of divergence of these isolates from prototype sequences of other subtypes was not different from intra-

subtype differences, strongly suggesting that the sequence variability in the HVR1 is independent of HCV genotype.

4.5 Variations in HCV antibody levels post-transplantation

Table 4.2 depicts values for antibody levels as determined by the ABBOTT anti-HCV 2nd generation kit. Values are determined for liver transplant recipients A1, C1, C2, F1, L1, M1 and M2, as well as candidates for OLT A2, B1 and C3, and asymptomatic blood donors Dix, Ad and Wad. All determinations are given as ratios of sample absorbance reading to cut-off value. All assays were performed by G. Pearson.

Table 4.4: *Absorbance readings for HCV. Dotted lines demarcate pre- from post-OLT readings.*

Patient	Sample	Days post-OLT	OD/cut-off ratio
A1	A1B1	-8	0.678
	A1B2	-3	0.897
	A1A1	4	0.383
	A1A2	12	0.505
	A1A3	18	0.380
	A1A4	25	0.429
	A1A5	32	0.550
	A1A6	42	0.608
	A1A7	49	0.535
	A1A8	56	0.574
	A1A9	70	0.696
	A1A10	74	0.608
	A1A11	98	0.620
	A1A12	109	0.729
C1	C1B1	-24	5.210
	C1A1	8	4.529
	C1A2	14	4.483
	C1A3	38	4.319
	C1A4	45	4.125
	C1A5	83	5.532
	C1A6	86	5.283

	C1A7	218	6.626
C2	C2B1	-18	5.942
	C2B2	-10	6.328
	C2A1	3	6.161
	C2A2	10	> 6.687
	C2A3	20	> 6.687
	C2A4	27	6.371
	C2A5	41	5.967
	C2A6	52	6.562
	C2A7	55	6.258
	C2A8	69	6.316
	C2A9	77	> 6.687
	C2A10	80	> 6.687
	C2A11	83	> 6.687
F1	F1A3	52	4.058
	F1A5	120	6.134
	F1A6	144	6.067
	F1A7	234	4.514
L1	L1B1	-17	> 6.687
	L1A1	8	> 6.687
	L1A2	36	6.450
	L1A3	64	6.289
	L1A4	88	5.796
M1	M1B2	-24	3.763
	M1B3	-15	3.769
	M1B4	-9	3.702
	M1A1	4	1.714
	M1A2	12	1.334
	M1A3	18	1.641
	M1A4	25	1.918
	M1A5	32	2.660
	M1A6	42	2.419
	M1A7	49	2.836
	M1A8	94	3.562

M2	M2A1	3	0.927
	M2A2	10	1.116
	M2A3	14	0.814
	M2A4	31	0.948
A2	A2-X	NTx*	5.456
B1	B1-X	ND*	1.796
C3	C3-X	ND	5.599
Ad		NA*	6.06
Wad		NA	>5.54
Dix	D1N2	NA	> 6.687

* NTx = not transplanted; ND = not done; NA = not applicable

The correlation between absorbance (OD) readings and antibody concentrations was explained in chapter 2 (section 2.2).

These results helped to explain some of the puzzling PCR results from section 4.1.1. A1 was antibody-negative both pre- and post-OLT, for all samples tested, a finding which is consistent with PCR results showing no evidence of HCV viraemia. However, analysis of peripheral blood leukocytes revealed the presence, confirmed by sequencing of the HVR1, of HCV in the lymphocytes.

Antibody levels were high for 4 patients C1, C2, F1 and L1. For the three patients (C1, C2 and L1) for whom both pre- and post-OLT samples were available, post-OLT antibody levels, as assessed by OD readings, were similar to pre-transplantation values. OD/cut-off ratios apparently fell post-OLT in one patient (M1), and tended to dip 1 - 3 months post-OLT in patients C1, C2 and L1. However, since no proper titration was performed, these differences were not considered significant.

As discussed in section 4.1.1 above, pre-transplantation samples for B1, C3 and M2 were repeatedly PCR-negative for HCV sequences. One possible reason is a level of viraemia too

low for detection by the method used, at least for both M2 and B1. Although unlikely, another possible reason is infection by an HCV strain sufficiently divergent to prevent annealing of the PCR primers, and which induces antibodies that do not react with the kit antigens. M2, who is of Egyptian origin, may have been infected with the predominant type 4 Egyptian strain. This would also explain the patient's borderline antibody response.

4.6 Spurious sequences amplified with HCV *E1/E2* primers

Primer 3766 (inner sense primer, see section 2.6.1, table 2.1) was used as a sequencing primer to determine the HCV genotype in all samples. In a significant number of cases, a non-HCV sequence was obtained from white blood cells. The spurious sequence, which was amplified from leukocytes of 4 patients (A1, C1, C2 and Dix) and from plasma of patient F1, is depicted in figure 4.9.

Figure 4.9: *Non-HCV sequence amplified from white blood cells of HCV-infected subjects*

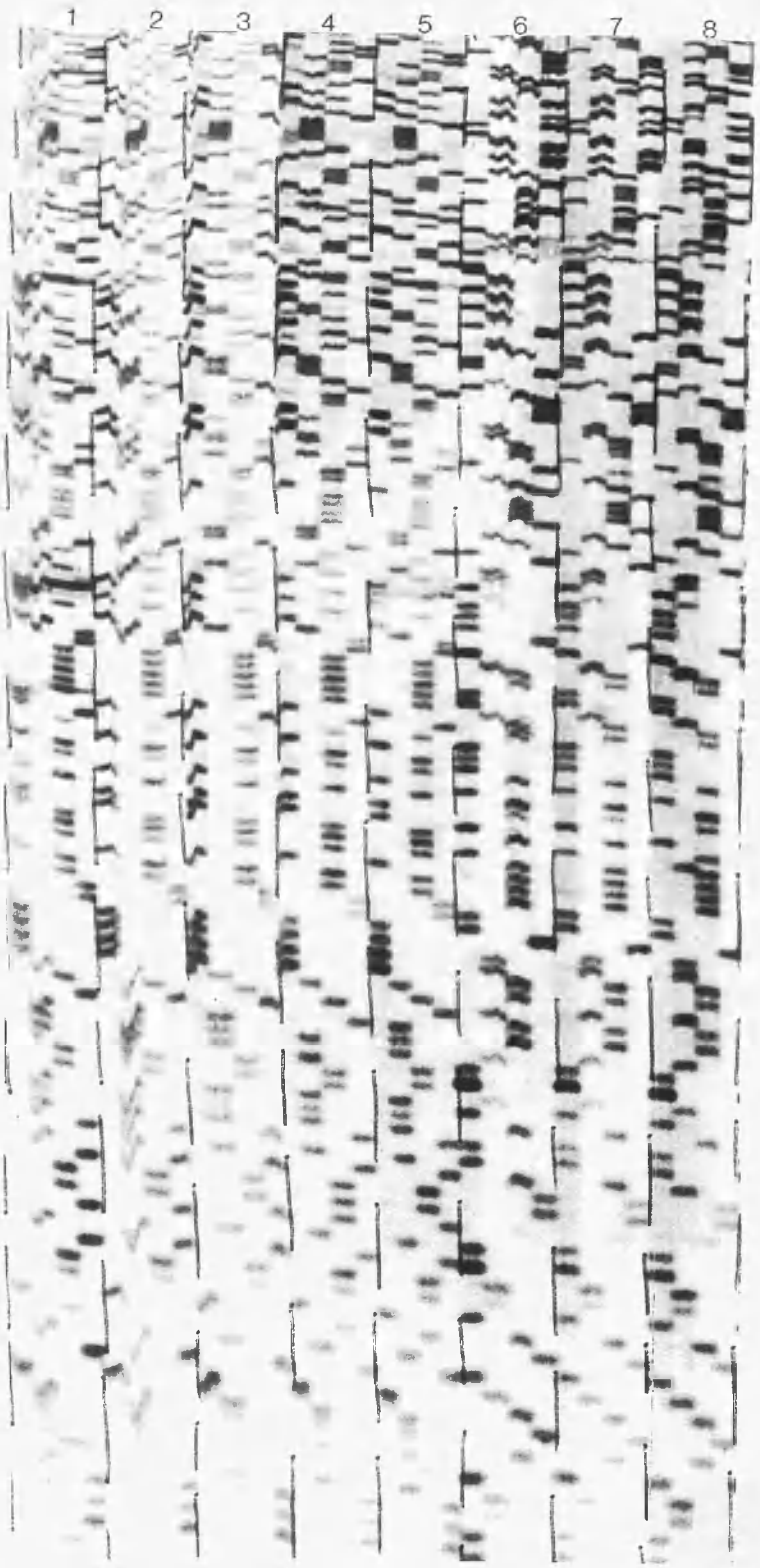
CACCCTTGAG	AGTTGAAGGG	TCGGAAACGC	CCTACGTGGA
CAGGACACCC	GGCCCAGCTT	TTAAGGTGGG	TTGTGGTGAT

This finding was surprising in view of the high specificity of the PCR primers, combined with the fact that nested PCR with two sets of primers was carried out, and that the PCR DNA band which gave rise to non-HCV sequences was indistinguishable by size from the HCV-specific band. Indeed, mixed HVR1 and spurious sequences were obtained from clones originating from the same amplicon (figure 4.10). The unusual sequence is most likely a part of the human genome because:-

- it was obtained mostly from leukocytes and not plasma, consistent with the human genomic content of the former;
- it was identical in 5 unrelated patients

Where the leukocytes had been sub-divided into lymphocyte- or polymorph-enriched populations, the non-HCV sequence was found exclusively in the polymorph fraction. Submission of the sequence to the various data banks revealed a lack of coding potential for this fragment.

Figure 4.10: *Sequences derived from Dix using E1/E2 primers showing random distribution of HCV and non-HCV sequences.*



Lanes were loaded in the order: TCGA. Nos. 1 - 5 = non-HCV sequences. Nos. 6 - 8 = HCV sequences

CHAPTER FIVE: HVR1 SEQUENCE VARIABILITY IN IMMUNOSUPPRESSED AND UNTREATED SUBJECTS

The HCV HVR1 is known to induce neutralising antibodies (Shimizu *et al.* 1994; Zibert *et al.* 1995). In order to assess the effect of host immune pressure on genomic variability in this region, serial samples obtained from immunosuppressed liver transplant recipients were analysed. Sets of at least 8 HVR1-containing cDNA clones were sequenced for each patient, and compared with similar data obtained from untreated, asymptomatic blood donors infected with HCV. For each set, the typing region (nucleotides 1330 - 1425) was also sequenced, to confirm that it remained essentially unchanged. Table 5.1 sums up the clinical histories of the seven individuals from whom serial samples were obtained. Asymptomatic individuals were designated patients 1 - 3, while the immunosuppressed patients were designated patients 4 -7.

Table 5.1: *Clinical histories of immunosuppressed and untreated HCV-infected subjects*

Subject		Sex	Age	Diagnosis[1]	HCV subtype[2]	Clinical outcome	Country of origin
Asymptomatic HCV hosts - untreated							
1	Wad	M	32	CAH	I/1a	alive	UK
2	Dix	M	44	CAH	IV/2b	alive	UK
3	Fra	F	37	CPH	II/1b	alive	UK
Immunosuppressed OLT recipients							
4	C1	M	56	CIRR/HCC	II/1b	dead	Italy
5	L2	M	48	CIRR	II/1b	dead	UK
6	L1	M	66	CIRR	II/1b	alive	Italy
7	L3	F	46	CIRR	I/1a	alive	Italy

1 CIRR - cirrhosis; HCC - hepatocellular carcinoma; CAH - chronic active hepatitis; CPH - chronic persistent hepatitis (defined on the basis of liver biopsies).

2 Roman and Arabic numerals refer, respectively to the Okamoto (Okamoto *et al.* 1992) and Simmonds (1995) classification systems - see section 1.5.3.

Figure 5.1: *Nucleotide (a) and deduced amino acid (b) sequences for HVR1 of asymptomatic blood donor Wad (patient 1)*

(a) HCV HVR1 NT SEQUENCES: Reference 1 (Wad)

	1								10								20															
	Thr	Thr	Tyr	Thr	Ser	Gly	Gly	Ser	Ala	Ala	Arg	Asp	Thr	Ala	Gly	Leu	Ala	Gly	Leu	Phe	Asn	Leu	Gly	Pro	Lys	Gln	Thr					
0 days:-	ACT	ACC	TAC	ACC	AGC	GGG	GGG	AGT	GCT	GCC	AGG	GAC	ACG	GCT	GGA	CTC	GCC	GGC	CTA	TTC	AAC	CTG	GGC	CCC	AAG	CAG	ACA					**No. of clones**
WAD101	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	4
WAD001	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	A..	...	...	...	...	...	...	..t	...	...	...	...	...	4
WAD114	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	A..	...	...	...	...	...	...	...	...	...	...	...	...	1
267 days:-																																
WAD203	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	7
WAD202	...	...	...	...	..t	...	..a	...	...	...	...	.C.	...	T..	.CG	...	.T.	..g	...	C..	.G.	.C.	...	G..	...	...	.AT					4

(b) HCV HVR1 AMINO ACID SEQUENCES: Reference 1 (Wad)

	+	1								1	0								2	0									
	T	T	Y	T	S	G	G	S	A	A	R	D	T	A	G	L	A	G	L	F	N	L	G	P	K	Q	T	
0 days:-																													**No. of clones**
WAD101	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4
WAD001/114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.	.	.	.	5
267 days:-																													
WAD203	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	7
WAD202	.	.	.	.	.	.	.	.	.	.	.	A	.	S	A	.	V	.	.	L	S	P	.	A	.	.	N		4

Dots (.) indicate nucleotide (a) or deduced amino acid (b) sequence identical to dominant sequence of first time point. Silent mutations shown in lower case. Nt. 1 corresponds to nt 1489 of HC-J8 (Okamoto et al., 1992).

5.1 HVR1 sequence variability in asymptomatic HCV-infected individuals

Patient 1 (Wad): Three HVR1 genomic variants were detected in the first plasma sample from patient 1 (Wad). As depicted in figure 5.1, these species coexisted in a ratio of 4:4:1, represented by WAD101, -001 and -114, respectively. All three were similar (maximum nucleotide sequence diversity = 2.5%). The second sample was obtained 267 days after the first. At this time point, two variants were present, in a ratio of 7:4 (represented by WAD203 and 202, respectively). One (WAD203) was identical to WAD101, but the other (WAD202) was extensively mutated, with 14 base substitutions in the 81-nucleotide HVR1. Most (9/14 or 64%) of the changes were in the first or second codon positions, leading to the replacement of 9 of the 27 deduced amino acids. Furthermore, most (6/9) of the amino acid changes were non-conservative. Nucleotide sequence diversity in this sample reached 17.3%. The sequences corresponding to WAD001 and WAD114 were undetectable at time point 2.

Patient 2 (Dix): The HVR1 quasispecies population in untreated patient 2 was determined at three time points, the second and third time points being 457 and 719 days after the first. Results are depicted in figure 5.2. At the first time point, two genomic variants were observed. The minor sequence (represented by PDN113) differed from the first (PDN111) by 11 nucleotides (13.6%), leading to a deduced amino acid sequence diversity of 7/27, or 26%. The quasispecies population changed considerably in the 14 months between the first and second time points. By time point 3, even greater diversity was found (up to 21% nucleotide difference), with the number of unique HVR1 species rising from 4 to 10.

Nucleotide sequence diversity between PDN113 and the original dominant sequence PDN111 occurred at HVR1 positions 25, 26, 27, 37, 38, 48, 49, 55, 62, 66 and 81. The mutations in 6 of these positions (26, 38, 48, 49, 55 and 62) were retained in 11 of 14 clones of time point two, implying the same lineage for both sets of sequences despite the 457-day interval. The lineage apparently persisted until time point 3, as most of these mutations were maintained in a significant number (6/11) of clones obtained at the final time point.

Figure 5.2: *Nucleotide (a) and deduced amino acid (b) sequences for HVR1 of asymptomatic blood donor Dix (patient 2)*

(a) HCV HVR1 NT SEQUENCES: Reference 2 (Dix)

	Thr	Thr	Tyr	Thr	Thr	Gly	Ala	Gln	Met	Gly	Arg	Gly	Ile	Thr	Gly	Phe	Ser	Asn	Leu	Phe	Asn	Leu	Gly	Ser	Gln	Gln	Lys	No. of clones
Position	1									10										20								
0 days:-																												
PDN111	ACA	ACC	TAT	ACC	ACC	GGA	GCA	CAA	ATG	GGC	AGA	GGC	ATT	ACG	GGG	TTT	TCT	AAC	CTC	TTC	AAC	TTG	GGT	TCC	CAG	CAG	AAG	
PDN111	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	3
PDN113	...	...	...	...	...	...	...	...	TCA	...	...	...	GC.	...	...	..G	G..	...	T..	...	.G.	..a	...	...	...	...	..C	2
457 days:-																												
PDN268	G..	...	...	.T.	...	...	...	...	.C.	...	...	...	.C.	G..	.CA	..G	G..	.G.	T..	..A	.G.	...	...	C..	...	...	...	11
PDN265	...	...	...	...	...	...	...	...	...	...	CA.	...	...	...	...	...	...	..G	...	...	...	A..	...	...	...	...	...	1
PDN266	.G.	...	...	...	...	...	...	...	.C.	...	.AC	...	GC.	G..	A..	..G	G..	.G.	T..	...	.G.	...	...	...	...	...	...	1
PDN267	G..	...	...	.T.	...	...	...	...	.C.	...	...	...	.C.	G..	.CC	..G	G..	.G.	T..	..A	.G.	...	...	C..	...	...	...	1
719 days:-																												
PDN301	...	...	C..	.T.	...	...	...	...	.C.	...	...	...	.C.	G..	.C.	..G	G..	...	T..	..A	.G.	...	...	C..	...	...	...	2
PDN302	...	...	...	...	...	...	...	...	GC.	...	...	A..	.C.	TG.	.C.	..c	G..	...	T..	...	.G.	CC.	...	G..	...	...	...	1
PDN303	...	...	...	...	...	...	...	...	GC.	...	...	A..	.C.	TG.	.C.	...	G..	...	T..	...	.G.	CC.	...	G..	...	...	...	1
PDN304	...	...	...	...	...	...	...	...	GC.	...	..C	A..	.C.	TT.	.C.	..G	G..	..G	T..	...	.G.	...	...	G..	...	...	../	1
PDN305	...	...	...	...	...	...	...	...	...	...	.A.	.C.	...	...	...	...	...	...	...	...	..A	...	...	...	...	...	...	1
PDN307	...	...	...	...	...	...	...	...	...	...	G..	...	GC.	.T.	...	...	...	...	...	...	.G.	...	...	...	...	...	...	1
PDN316	...	...	...	...	...	...	...	...	GC.	...	...	A..	.C.	TG.	.C.	..G	G..	...	A..	...	.G.	...	...	...	...	...	...	1
!PDN318	...	...	...	...	...	...	...	...	...	...	...	...	GC.	.T.	...	...	...	...	...	...	...	...	...	...	...	AC.	...	1
PDN319	...	...	...	...	...	...	...	...	GCA	...	...	A..	.C.	TG.	.C.	..G	G..	...	A..	...	.G.	CC.	...	G.T	...	...	...	1
PDN320	...	...	...	...	...	...	...	...	...	...	CA.	.C.	GC.	GG.	...	..c	G..	...	...	...	.G.	...	...	...	...	...	...	1

(b) HCV HVR1 DEDUCED AMINO ACID SEQUENCES: Reference 1 (Wad)

	+																											
Position	1									10										20								
PDN111	T	T	Y	T	T	G	A	Q	M	G	R	G	I	T	G	F	S	N	L	F	N	L	G	S	Q	Q	K	No. of clones
0 days:-																												
PDN111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3
PDN113	.	.	.	.	.	.	.	.	S	.	.	.	A	.	.	L	A	.	F	.	S	.	.	.	.	.	N	2
457 days:-																												
PDN268	A	.	.	I	.	.	.	.	T	.	.	.	T	A	A	L	A	S	F	L	S	.	.	P	.	.	.	11
PDN265	.	.	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	K	.	.	.	M	.	.	.	.	.	1
PDN266	R	.	.	.	.	.	.	.	T	.	N	.	A	A	R	L	A	S	F	.	S	.	.	.	.	.	.	1
PDN267	A	.	.	I	.	.	.	.	T	.	.	.	T	A	A	L	A	S	F	L	S	.	.	P	.	.	.	1
719 days:-																												
PDN301	.	.	H	I	.	.	.	.	T	.	.	.	T	A	A	L	A	.	F	L	S	.	.	P	.	.	.	2
PDN302	.	.	.	.	.	.	.	.	A	.	.	S	T	W	A	.	A	.	F	.	S	P	.	A	.	.	.	1
PDN303	.	.	.	.	.	.	.	.	A	.	.	S	T	W	A	.	A	.	F	.	S	P	.	A	.	.	.	1
PDN304	.	.	.	.	.	.	.	.	A	.	S	S	T	L	A	L	A	K	F	.	S	.	.	A	.	.	.	1
PDN305	.	.	.	.	.	.	.	.	.	.	K	A	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	1
PDN307	.	.	.	.	.	.	.	.	.	.	G	.	A	M	.	.	.	.	.	.	S	.	.	.	.	.	.	1
PDN316	.	.	.	.	.	.	.	.	A	.	.	S	T	W	A	L	A	.	I	.	S	.	.	.	.	.	.	1
!PDN318	.	.	.	.	.	.	.	.	.	.	.	.	A	M	.	.	.	.	.	.	.	.	.	.	.	T	.	1
PDN319	.	.	.	.	.	.	.	.	A	.	.	S	T	W	A	L	A	.	I	.	S	P	.	A	.	.	.	1
PDN320	.	.	.	.	.	.	.	.	.	.	Q	A	A	G	.	.	A	.	.	.	S	.	.	.	.	.	.	1

Dots (.) indicate nucleotide (a) or deduced amino acid (b) sequence identical to dominant sequence of first time point. Silent mutations shown in lower case. Nt. 1 corresponds to nt 1489 of HC-J8 (Okamoto et al., 1992).

Patient 3 (Fra): Figure 5.3 shows the HVR1 quasispecies distributions in three samples collected from patient 3 during a three-month period. Dynamic fluctuations were observed - none of the quasispecies detected in the final sample had been present at the initial sample time point. Interestingly, the number of molecular species detected rose from 3 to 8 in the 4-week period between the first and second samples. By the third time point (approximately two months after the second), the number of species had fallen to 2. Considerable nucleotide and amino acid sequence diversity was observed at all time points, similar to that found in patient 2, although in patient 3 the samples were obtained over a considerably shorter period.

In the initial sample, the dominant sequence (denoted by FRA104) represented 9 of 11 clones, a tenth clone (FRA101) differing by only one nucleotide. The other minor variant (FRA110) was considerably different, with a nucleotide sequence divergence of 22% (18 of 81 nucleotides) from the dominant sequence. Replacement mutations in this sequence corresponded to 13/27 (48%) amino acid changes, only 4 of them conservative. A decline in the relative proportion of the FRA104 sequence (which made up 82% of the initial population) was noted at the intermediate time point, when it comprised 25% of the population sequenced. Only 5 of 12 clones were still clearly related to the initially dominant FRA104 quasispecies at time point 2. All five had none or a single base substitution, and a deduced amino acid sequence identical to the FRA104 sequence. The remaining 7 clones, which sometimes differed substantially from each other, appeared nevertheless to be mutants of FRA110, a minor variant of the first time point. Index mutations at nucleotide positions 1, 22, 25, 37, 43, 52, 58, 62, 63, 65 and 79 were present in all. Furthermore, they shared 6 amino acid replacements, as well as up to 8 additional amino acid changes.

In the third sample, collected two months after the second, each of the two molecular species (symbolised by FRA301 and 302) was clearly related to either FRA104 or FRA110, the major and minor variants, respectively, of the first time point. However, the lineage of FRA110 (FRA301 and others) had become dominant, accounting for 7 of the 12 clones sequenced. The "FRA302 series" showed amino acid sequences identical with that of FRA104.

Figure 5.3: *Nucleotide (a) and deduced amino acid sequences (b) for HVR1 of asymptomatic blood donor Fra (patient 3)*

(a) HCV HVR1 NT SEQUENCES: Reference 3 (Fra)

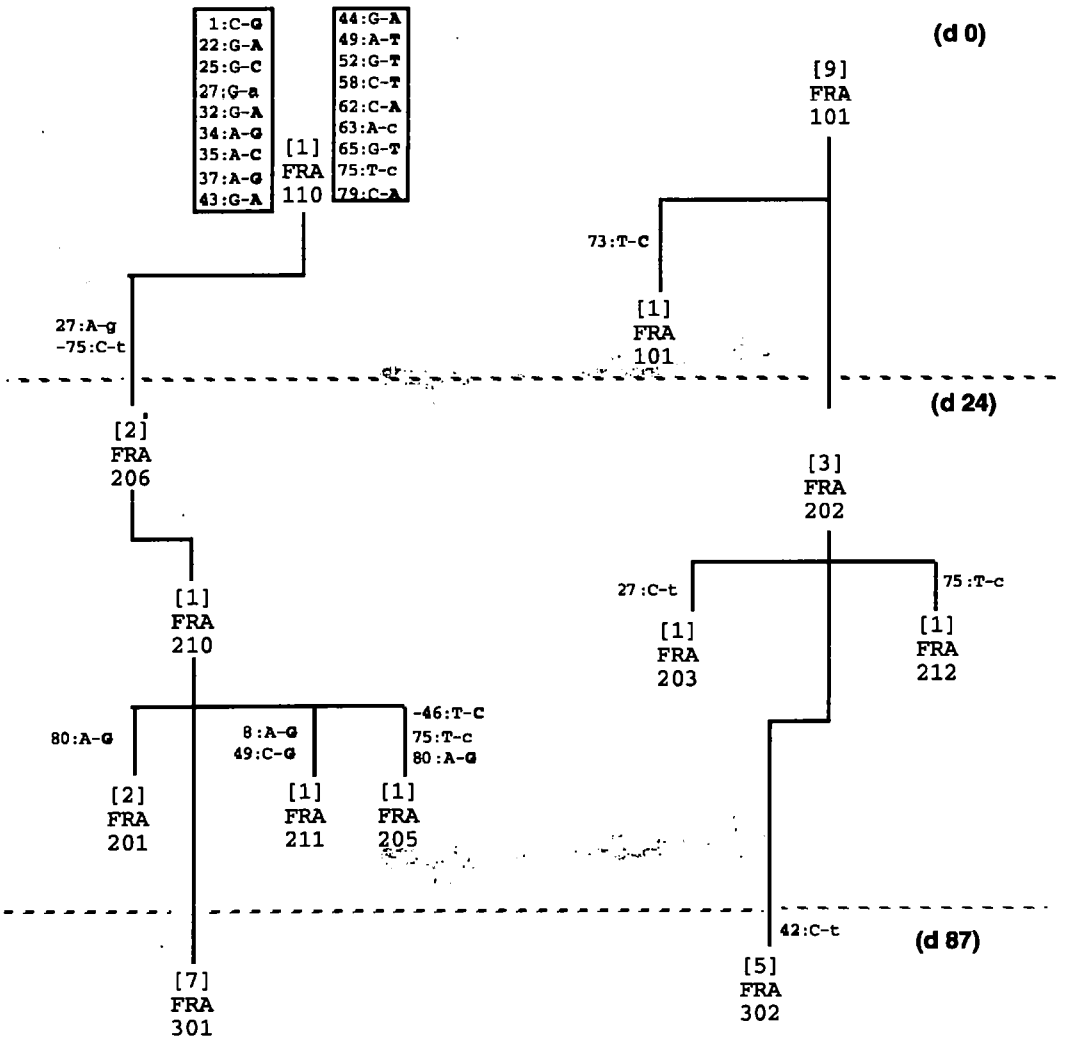
	Gln	Thr	His	Thr	Thr	Gly	Gly	Ala	Val	Ala	Arg	Asn	Thr	Tyr	Gly	Leu	Thr	Ala	Leu	Leu	Thr	Arg	Gly	Pro	Ser	Gln	Gln	No. of clones
	1									10										20								
0 days:-																												
	CAA	ACC	CAC	ACG	ACA	GGG	GGG	GCA	GTC	GCC	CGC	AAC	ACC	TAC	GGG	CTT	ACG	GCC	CTC	CTC	ACA	CGT	GGG	CCG	TCT	CAG	CAA	
FRA104	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	9
FRA101	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	C..	...	...	1
FRA110	G..	...	...	...	...	...	...	A..	C.A	...	.A.	GC.	G..	...	AA.	...	T..	T..	...	T..	.AC	.T.	...	...	..c	...	A..	1
24 days:-																												
FRA202	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	3
FRA201	G..	...	...	GT.	...	...	...	A..	CAA	.G.	...	.T.	GT.	A..	A.A	T..	...	T..	T..	T..	.AC	.T.	...	...	...	...	AG.	2
FRA206	G..	...	...	...	...	...	...	A..	C.G	...	.A.	GC.	G..	...	AA.	...	T..	T..	...	T..	.AC	.T.	...	...	...	...	A..	2
FRA211	G..	...	.G.	GT.	...	...	...	A..	CAA	.G.	...	.T.	GT.	A..	A.A	T..	G..	T..	T..	T..	.AC	.T.	...	...	...	...	A..	1
FRA203	...	...	...	...	...	...	...	...	..t	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1
FRA210	G..	...	...	GT.	...	...	...	A..	CAA	.G.	...	.T.	GT.	A..	A.A	T..	...	T..	T..	T..	.AC	.T.	...	...	...	...	A..	1
FRA212	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	..c	...	...	1
FRA205	G..	...	...	GT.	...	...	...	A..	CAA	.G.	...	.T.	GT.	A..	A.A	...	...	T..	T..	T..	.AC	.T.	...	...	...	...	AG.	1
87 days:-																												
FRA301	G..	...	...	GT.	...	...	...	A..	CAA	.G.	...	.T.	GT.	A..	A.A	T..	...	T..	T..	T..	.AC	.T.	...	...	...	...	A..	7
FRA302	...	...	...	...	...	...	...	...	...	...	...	...	...	..t	...	...	...	...	...	...	...	...	...	...	...	...	...	5

(b) HCV HVR1 DEDUCED AMINO ACID SEQUENCES: Reference 3 (Fra)

	Q	T	H	T	T	G	G	A	V	A	R	N	T	Y	G	L	T	A	L	L	T	R	G	P	S	Q	Q	No. of clones
	+ 1									1 0										2 0								
0 days:-																												
FRA104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	9
FRA101	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	.	1
FRA110	E	.	.	.	.	.	.	T	L	.	H	A	A	.	K	.	S	S	.	F	N	L	.	.	.	.	K	1
24 days:-																												
FRA202	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3
FRA201	E	.	.	V	.	.	.	T	Q	G	.	I	V	N	R	F	.	S	F	F	N	L	.	.	.	.	R	2
FRA206	E	.	.	.	.	.	.	T	L	.	H	A	A	.	K	.	S	S	.	F	N	L	.	.	.	.	K	2
FRA211	E	.	R	V	.	.	.	T	Q	G	.	I	V	N	R	F	A	S	F	F	N	L	.	.	.	.	K	1
FRA203	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
FRA210	E	.	.	V	.	.	.	T	Q	G	.	I	V	N	R	F	.	S	F	F	N	L	.	.	.	.	K	1
FRA212	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
FRA205	E	.	.	V	.	.	.	T	Q	G	.	I	V	N	R	.	.	S	F	F	N	L	.	.	.	.	R	1
87 days:-																												
FRA301	E	.	.	V	.	.	.	T	Q	G	.	I	V	N	R	F	.	S	F	F	N	L	.	.	.	.	K	7
FRA302	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5

Dots (.) indicate nucleotide (a) or deduced amino acid (b) sequence identical to dominant sequence of first time point. Silent mutations shown in lower case. Nt. 1 corresponds to nt 1489 of HC-J8 (Okamoto et al., 1992).

Figure 5.4: *Evolutionary tree of HVR1 sequences in asymptomatic blood donor Fra (patient3)*



Nature and positions of nucleotide sequence divergence from major quasispecies (FRA101 sequence) shown for three time points. Replacement mutations indicated by capitalized, emboldened letters, silent mutations by small letters. E.g., "39:T-c" = silent T to C mutation at position 39 of HVR1. In reverse mutation positions, the minus (-) sign precedes the HVR1 nucleotide position. Boxed mutations (first time point only) represent positions of nucleotide sequence divergence from dominant quasispecies.

Figure 5.4 shows the deduced evolutionary relationships between the various clones from Fra. It can be seen that underlying the numerous mutations is a certain degree of consistency. The sequences of the first time point, as well as all mutants observed thereafter, may be roughly divided, on the basis of nucleotide sequence similarities, into two groups or "families",

corresponding to either FRA104 or FRA110. Thus, of the 8 variants detected at time point 2, the FRA202, -203 and -212 entities, representing 5/12 clones, or 42% of the quasispecies population, are identical with, or unambiguously related to, the FRA101 "family". The remaining 5 species, represented by FRA201, -206, -211, -210 and -205, are evidently derived from FRA110, despite a nucleotide and amino acid sequence divergence reaching 31% and 67%, respectively. Likewise, the FRA301 set of clones for time point 3, which constitute a single variant representing 7 (52%) of 12 clones, are related to FRA110, while the FRA302 variants have the FRA104 progenitor.

5.2 HVR1 sequence variability in HCV-infected liver transplant recipients

HCV HVR1 sequences were obtained from four patients, who each received a liver allograft for end-stage cirrhosis and were followed up for 3 to 15 months. Sequences were obtained within the month preceding and by two months following OLT for 3 patients (C1, L1 and L3). For an additional liver transplant recipient, L2, sequences were obtained corresponding to five time points spanning one year post-OLT.

***Patient 4 (C1)*:** A marked lack of nucleotide sequence diversity was evident in sequences obtained from immunosuppressed OLT recipient C1, which was even more pronounced at the amino acid level (Fig 5.5). Although 6 molecular variants were detected pre-OLT, they differed from each other and from the major sequence by no more than 5 nucleotides, and 8 of 9 points of nucleotide sequence diversity gave the same deduced amino acid as the major sequence. The number of variants detected had fallen to two by 8 days post-OLT, and a minor pre-OLT species, carrying a replacement mutation (Ile to Leu) had become the predominant post-OLT strain.

Prior to transplantation, 6 variants were detected among 13 clones derived from recipient C1. Distributed in the ratio 7:2:1:1:1:1, they are represented by SCB002, -004, -010, -012, -013 and -014 in figure 5.5. As observed with the asymptomatic HCV hosts, there appeared to be two "families" of HVR1 sequences. The first comprised the dominant sequence (SCB002)

Figure 5.5: *Nucleotide (a) and deduced amino acid sequences (b) for HVR1 of liver transplant recipient C1 (patient 4)*

(a) HCV HVR1 NT SEQUENCES: C1

```
          1           5              10              15              20              25
          |           |              |               |               |               |
          Thr Thr Tyr Val Thr Gly Gly Ser Ala Ala Arg Arg Thr Gln His Val Thr Ser Ile Phe Ser Phe Gly Ser Ser Gln Asn     No. of
Pre-OLT:-                                                                                                              clones
SCB002    ACC ACC TAC GTG ACG GGG GGG TCA GCA GCC CGC CGA ACT CAG CAT GTG ACA TCC ATC TTT TCA TTT GGG TCG TCT CAG AAC       7
SCB004    ..t ... ... ... ... ... ... ... ..g ... ... ... ..c ... ... ... ... ... ... ... ... ... ... ... ... ...       2
SCB010    ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ..a ... ..t C.. ... ... ..c ... ... ... ... ...       1
SCB012    ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ... ... ..c ... ... ... ... ...       1
SCB013    ... ... ... ... ..a ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ..c ... ..c ... ... ... ... ...       1
SCB014    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..g ... ... ... ... ... ...       1

8 days post-OLT:-
PCA115    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ..c ... ..c ... ... ... ... ...      10
PCA116    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ... ... ..c ... ... ... ... ...       1

86 days post-OLT:-
PCA605    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ... ... ..c ... ... ... ... ...       8

218 days post-OLT:-
PCA702    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ... ... ..c ... ... ... ... ...       9
PCA704    ... ... ..t ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ... ... ..c ... ... ... ... ...       1
PCA707    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... T.. ... ... ..t C.. ... ... ..c ... ... ... ... ...       1

329 days post-OLT:-
PCA801    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ... ... ..c ... ... ... ... ...       7
PCA804    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ... C.. ... ... ..c ... ... ... ... ...       2
PCA809    ... ... ... ... ... ... ... ... ... ..g ... ... ... ..c ... ... ... ... ..t C.. ... ... ..c ... ... ... ... ...       2
PCA808    ... ... ... ... ... ... ... ... ... ... ... ... ... ..c ... ... ... ... ..t C.. ... ..g ..c ... ... ... ... ...       1
```

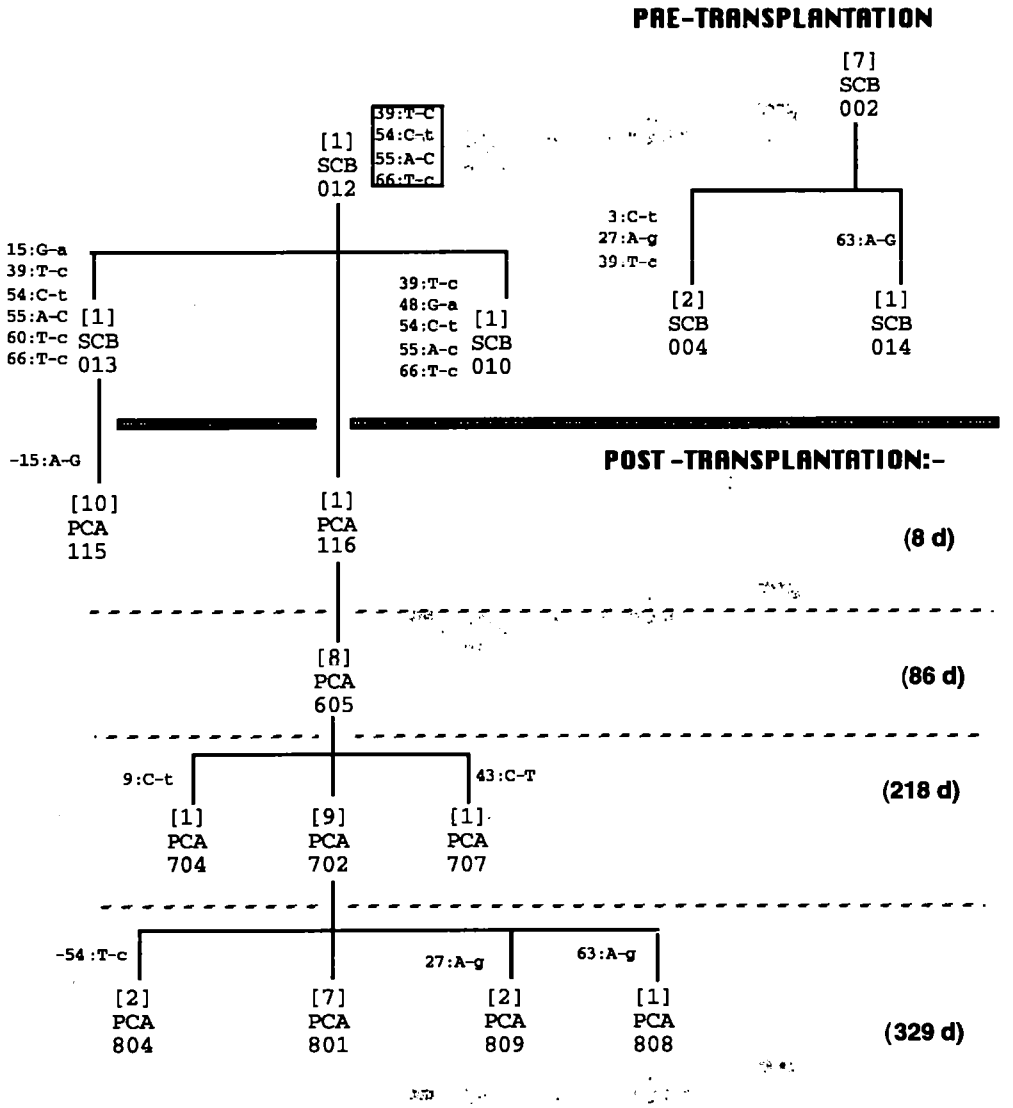
(b) HCV HVR1 DEDUCED AMINO ACID SEQUENCES: C1

```
          + 1                   10                  20                                                    No. of
Pre-OLT:-|                      |                   |                                                     clones
SCB002/   T T Y V T G G S A A R R T Q H V T S I F S F G S S Q N
004/014   . . . . . . . . . . . . . . . . . . . . . . . . . . .                                            10
SCB010/0
012/013   . . . . . . . . . . . . . . . . . . . L . . . . . . .                                             3


8 days post-OLT:-
PCA115/16 . . . . . . . . . . . . . . . . . . . L . . . . . . .                                            11


86 days post-OLT:-
PCA605    . . . . . . . . . . . . . . . . . . . L . . . . . . .                                             8

218 days post-OLT:-
PCA702/4  . . . . . . . . . . . . . . . . . . . L . . . . . . .                                            10
PCA707    . . . . . . . . . . . . . . . . . Y . L . . . . . . .                                             1


329 days post-OLT:-
PCA801/04
 09/08    . . . . . . . . . . . . . . . . . . . L . . . . . . .                                            12
```

Dots (.) indicate nucleotide (a) or deduced amino acid (b) sequence identical to dominant sequence of first time point. Silent mutations shown in lower case. Nt. 1 corresponds to nt 1489 of HC-J8 (Okamoto et al., 1992).

and two minor species (SCB004 and -014), and accounted for a total of 10 clones, or 77% of the 13 clones sequenced. The second "family" shared divergence from SCB002 at positions 39, 54, 55 and 66. Only one divergent nucleotide (C for A at position 55) led to an amino acid replacement - Leu for Ile at amino acid 19 of the HVR1.

Figure 5.6: *Evolutionary tree of HVR1 sequences in patient C1*



Nature and positions of nucleotide sequence divergence from major species (SCB002 sequence) shown for pre-OLT time point, as well as for mutants seen thereafter. Silent mutations indicated by small letters, replacement mutations capitalized and emboldened. In reverse mutation positions, the minus (-) sign precedes the HVR1 nucleotide position. Boxed mutations (first time point only) represent points of nucleotide sequence divergence from major pre-OLT variant.

By 8 days post-OLT, the minor "family" of variants had become exclusive, remaining so throughout the approximately one year of follow-up. The highest value for HVR1 nucleotide sequence divergence was 8.6% for the first sample, and never exceeded 2.5% in any subsequent sample. Nucleotide sequence diversity was reduced to 0% in the population of clones obtained 86 days post-OLT, and maximal amino acid sequence diversity for any time point was 4%. This was in marked contrast with HVR1 sequences from patient 3 (Fra), in whom maximal nucleotide sequence diversity at the 3 time points (1 - 3) sampled were 23.5%, 30.9% and 28.4%, respectively. The corresponding values for amino acid sequence divergence were 58%, 67% and 59%. As with patient 1 (Fra), the sequencing results can be presented as an evolutionary tree, shown in Figure 5.6.

Patient 5 (L2): As with patient C1, HVR1 mutations observed in patient L2 during a year of follow-up were few and predominantly silent (see fig. 5.7). The dominant sequence remained essentially unchanged for all time points examined.

5 serum samples, collected 1, 3, 6, 12 and 15 months post-transplantation, were analysed. A pre-OLT sample could not be obtained for this patient. In the first sample, obtained 31 days post-OLT, the 5 detected HVR1 variants were arranged in the ratio 8:2:1:1:1, represented, respectively, by XL2A05, -A04, -A07, -A13 and -A14 (figure 5.7). 11 of the 13 clones fell into the same "family", 8 of them constituting the dominant species. A minor "family" of two clones had in common nucleotide mutations at positions 3, 6, 10, 11, 22, 25, 26, 30, 35 and 40 (also see figure 5.8).

At three months post-transplantation, the XL2A05 sequence remained dominant (see XL2C02) in a species partitioning of 6:5:1 representing XL2C02, -C03 and -C16, respectively. The latter two, although mutated, were clearly derived from the same "family" as XL2A05. In samples taken at 6 and 12 months post-OLT, the XL2A05 sequence persisted as the dominant species (77% and 93% of the respective populations), with a maximum nucleotide sequence diversity of only 1.2%. Amino acid sequence diversity remained 0%, as for the preceding time point. The last time point corresponded to two weeks following a second

Figure 5.7: *Nucleotide (a) and deduced amino acid (b) fsequences for HVR1 of liver transplant recipient L2 (patient 5)*

(a) HCV HVR1 NT SEQUENCES: L2

```
                1                                  10                                  20
                |                                  |                                   |
                Gln Thr Tyr Val Thr Gly Gly Ala Ser Ala Arg Asn Thr His Gly Ile Ala Ser Leu Phe Ala Phe Gly Pro Ala Gln Lys      No. of
31 days post-OLT1:-                                                                                                          clones
                CAA ACT TAT GTG ACA GGG GGG GCG TCA GCC AGA AAC ACC CAC GGG ATA GCG TCC CTC TTT GCC TTT GGA CCG GCT CAG AAA
XL2A05          ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...        8
XL2A04          ... ..c ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...        2
XL2A07          ..T ..c ... AC. ... ... ... T.. CA. ..t ... .C. ... T.. ... ... ... ... ... ... ... ... ... ... ... ...        1
XL2A13          ... ... ... ... ... ... ... ... ... ... ... ... G.. ... ... ... ... ... ... ... ... ... ... ... ... ...        1
XL2A14          ..T ..c ... AC. ... ... ... T.. CA. ..t ... .C. ... T.. ... ... ... ... ... ..c ... CC. ..g ... T.. ... ...        1

97 days post-OLT1:-
XL2C02          ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...        6
XL2C03          ... ..c ... ... ... ... ... ... ..g ... ... ... ..t ... ... ... ... ... ..c ... ... ... ... ... ... ...        5
XL2C16          ... ..c ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..t ... ... ... ... ... ...        1

180 days post-OLT1:-
XL2F02          ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...       10
XL2F05          ... ... ... ... ... ... ... ... ... ... ..g ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...        3

334 days post-OLT1:-
XL2Y03          ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...       13
XL2Y10          ... ... ... ... ... ... ... ... ..t ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...        1

435 days post-OLT1 (17 days post-OLT2):-
PL2102          ... ... ... ... ... ... ... ... ... ... ..g ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...       11
PL2109          ... ... ... ... ... ... ... ... ... ... ..g .G. ... ... ... ... ... ... ... ... ... ... ... ... ... ...        1
PL2115          ... ... ... ... ... ... ... ... ... ... ..g ... ... ... ... ... ... ... ... ... .C. ... ... ... ... ...        1
PL2118          ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... C.. ... ... ... ... ... ...        1
```

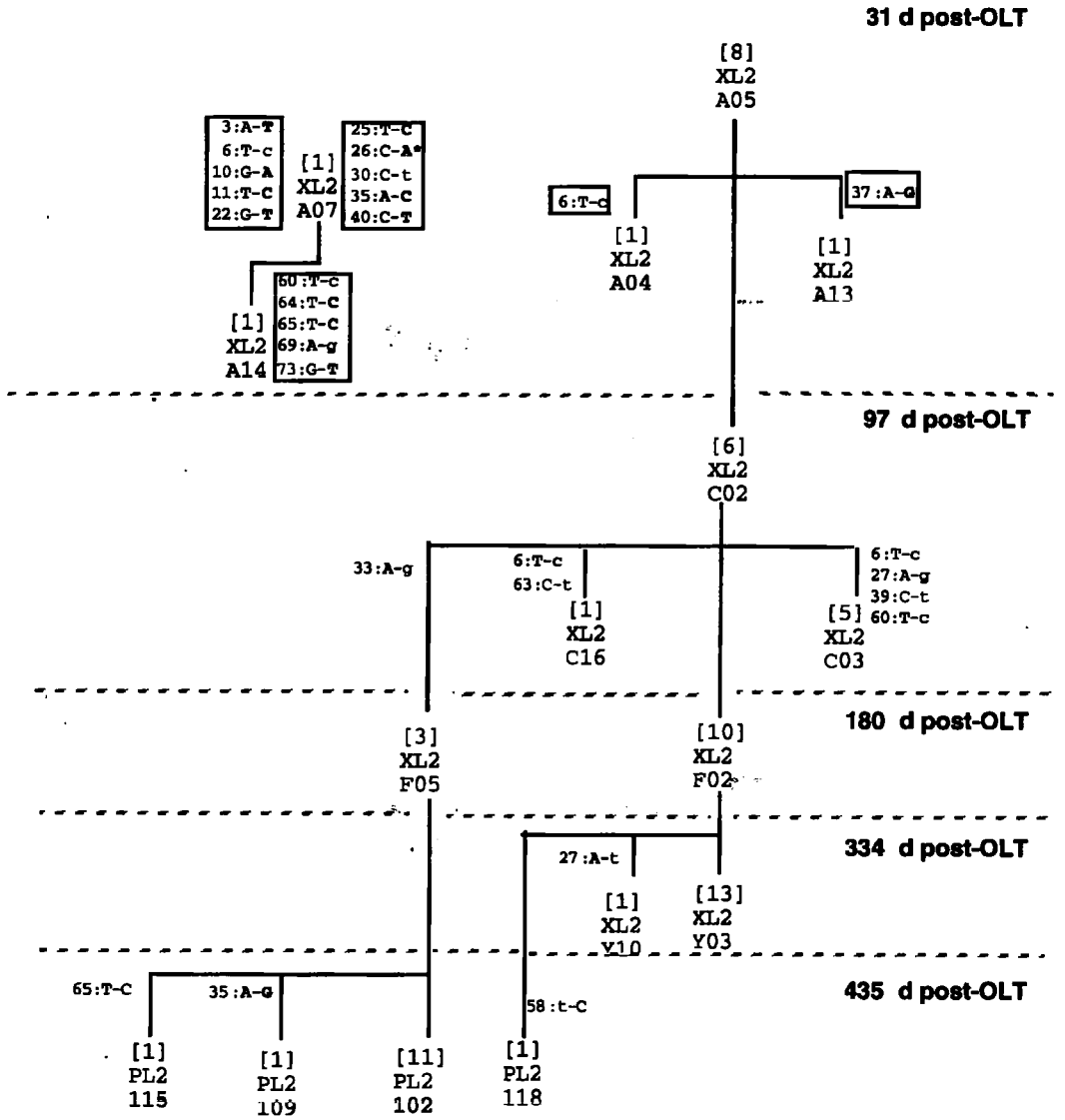
(b) HCV HVR1 DEDUCED AMINO ACID SEQUENCES: L2

```
              + 1                  10                  20
              |                    |                   |
               Q T Y V T G G A S A R N T H G I A S L F A F G P A Q K      No. of
31 days post-OLT1:-                                                     clones
XL2A05/04.    . . . . . . . . . . . . . . . . . . . . . . . . . .       10
XL2A07        H . . T . . . S Q . . T . Y . . . . . . . . . . . .        1
XL2A13        . . . . . . . . . . . . A . . . . . . . . . . . . /        1
XL2A14        H . . T . . . S Q . . T . Y . . . . . . . . P . . S . . /        1

97 days post-OLT1
XL2C02/03/
    16        . . . . . . . . . . . . . . . . . . . . . . . . . .       12

180 days post-OLT1
XL2F02/05.    . . . . . . . . . . . . . . . . . . . . . . . . . .       13

334 days post-OLT1
XL2Y03/10.    . . . . . . . . . . . . . . . . . . . . . . . . . .       14

435 days post-OLT1 (17 days post-OLT2)
PL2102        . . . . . . . . . . . . . . . . . . . . . . . . . .       11
PL2109        . . . . . . . . . . . S . . . . . . . . . . . . . .        1
PL2115        . . . . . . . . . . . . . . . . . . . . . S . . . .        1
PL2118        . . . . . . . . . . . . . . . . . . . . . L . . . .        1
```

Dots (.) indicate nucleotide (a) or deduced amino acid (b) sequence identical to dominant sequence of first time point. Silent mutations shown in lower case. Nt. 1 corresponds to nt 1489 of HC-J8 (Okamoto et al., 1992).

transplant, OLT2 (435 days post-OLT1). Nucleotide and deduced amino acid sequence diversity increased slightly to 2.5% and 3.7%, respectively, but the "XL2A05" sequence represented 78.5% of the sequenced clones.

Figure 5.8: *Evolutionary tree of HVR1 sequences in patient L2*



*Nature and positions of nucleotide sequence divergence from major species (XL2A05 sequence) shown for five post-OLT time points. Silent mutations indicated by small letters, replacement mutations capitalized and emboldened. * indicates mutation produces stop codon. In reverse mutation positions, the minus (-) sign precedes the HVR1 nucleotide position. Boxed mutations (first time point only) represent points of nucleotide sequence divergence from major variant at 31 days.*

Patient 6 (L1): Two samples were obtained from this patient, one pre- and one post-OLT. The 3 molecular species observed pre-transplantation are represented in figure 5.9 by SLB001, -003 and -015, found in the ratio 8:3:2, respectively. SLB003, which differed from SLB001 at one nucleotide position encoding a silent mutation, together with it comprised a single family representing 85% of the clones sequenced. The second minor variant detected (SLB015) accounted for a nucleotide sequence divergence of 9.9%. 7 of its 8 nucleotides which diverged from SLB001 were replacement mutations, leading to an amino acid sequence divergence of 6/27, or 22.2%.

The second sample from patient 6 was taken 3 months post-OLT (four months after the first sample). The number of HVR1 sequences detected had increased to 6, in the ratio 6:4:1:1:1:1, represented in fig. 5.8 by PLA402, -404, -424, -420, -403 and -406, respectively. However, maximal nucleotide sequence diversity had decreased slightly, from 9.9% pre- to 7.4% post-transplantation. All post-OLT variants had been undetectable in the pre-OLT samples. However, PLA402, -404, -424 and -420 were clearly mutants of SLB015, while PLA403 and -406 were probably derived from SLB003. Therefore, as in patient 4 (C1) and patient L3 (discussed below), a pre-OLT minor species became dominant after transplantation.

Patient 7 (L3): The 10 clones sequenced pre-OLT showed significant homogeneity of the HCV HVR1 population. Two variants were observed, in a ratio of 9:1, represented in figure 5.9 by PL3001 and -022, respectively. By 47 days post-OLT, the minor pre-OLT variant had become almost exclusive, constituting 11 of the 12 clones sequenced (represented by PL3101). A minor (8.3%) species (PL3107) was also detected at this time point, which may have arisen independently of PL3022. All 3 amino acid changes were unique, and two effected non-conservative changes, including a Ser to Pro alteration. Nucleotide and amino acid sequence diversities were comparable pre- and post-transplantation.

Figure 5.9: *Nucleotide (a) and deduced amino acid (b) sequences for HVR1 of liver transplant recipient L1 (patient 6)*

(a) HCV HVR1 NT SEQUENCES: L1

```
          1           5           10          15          20          25
          |           |           |           |           |           |
          Gln Thr His Thr Val Gly Gly Val Ala Gly Ser Thr Thr Ser Gly Phe Ala Ser Leu Phe Lys Phe Gly Pro Ser Gln Lys   No. of
Pre-OLT:                                                                                                              clones

33 days pre-OLT:

SLB001    TAA ACT CAC ACG GTA GGG GGG GTG GCA GGC TCC ACC ACC TCT GGG TTT GCG TCC CTC TTC AAG TTT GGG CCG TCT CAG AAA    8
SLB003    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..a ... ... ... ... ...    3
SLB015    ... ... ... ... ... ... ... ... .T. ... .T. T.. ... ... ... ... A.. ... ... ... .CA CG. ... ... ... ... ...    2

Post-OLT (88 days):-

PLA402    ... ... ... ... ... ... ... ... ... ... CT. ... ... ... ... ... A.. ... ... ... .GA C.. ... ... ... ... ...    6
PLA404    ... ... ... ... ..g ... ... ... ... ... CT. ... ... ... ... ... A.. ... ... ... .GA C.. ... ... ... ... ...    4
PLA424    ... ... ... ... .G. ... ... ... ... ... CT. ... ... ... ... ... A.. ... ... ... .GA C.. ... ... ... ... ...    1
PLA420    ... ... ... ... ..r ... ... ... ... ... ... ... ... ... ... ... A.. ... ... ... .GA C.. ... ... ... ... ...    1
PLA403    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..a C.. ... ... ... ... ...    1
PLA406    ... ... ... ... A.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..a C.. ... ... ... ... ...    1
```

(b) HCV HVR1 DEDUCED AMINO ACID SEQUENCES: L1

```
          1               10              20
Pre-OLT:  |               |               |                                   No. of clones
SLB001/   Q T H T V G G V A G S T T S G F A S L F K F G P S Q K
   003    . . . . . . . . . . . . . . . . . . . . . . . . . .        11
SLB015    . . . . . . . . V . F S . . . . T . . . T R . . . . .        2

Post-OLT (88 days):
PLA402/04 . . . . . . . . . . L . . . . . T . . . R L . . . . .       10
PLA424    . . . . G . . . . . L . . . . . T . . . R L . . . . .        1
PLA420    . . . . T . . . . . . . . . . . T . . . R L . . . . .        1
PLA403    . . . . . . . . . . . . . . . . . . . . L . . . . .        1
PLA406    . . . . I . . . . . . . . . . . . . . . L . . . . .        1
```

Dots (.) indicate nucleotide (a) or deduced amino acid (b) sequence identical to dominant sequence of first time point. Silent mutations shown in lower case. Nt. 1 corresponds to nt 1489 of HC-J8 (Okamoto et al., 1992).

Figure 5.10: *Nucleotide (a) and deduced amino acid (b) sequences for HVR1 of liver transplant recipient L3 (patient 7)*

(a) HCV HVR1 NT SEQUENCES: L3

```
        1                              10                  15                  20                  25
        |                              |                   |                   |                   |
        Gly Thr Tyr Thr Thr Gly Gly Ala Ala Gly Gln Thr Thr Ser Gly Leu Ala Gly Leu Phe Asn Ser Gly Ala Arg Gln His I     No. of
Pre-OLT:-                                                                                                            clones
              31                                    61                                    91
PL3001  GGC ACT TAC ACC ACC GGG GGA GCT GCC GGA CAG ACC ACA AGC GGG CTC GCC GGC CTC TTT AAT TCA GGG GCT AGA CAG CAC A      9
PL3022  ... ... ... ... ... ... ... ... ... ... ... ... ... C.. ... ... ... ... ... ... ... .CA ... ... ... ... ... ... .      1

Post-OLT (47 days):-
PL3101  ... ... ... ... ... ... ... ... ... ... ... ... ... C.. ... ... ... ... ... ... ... .CA ... ... ... ... ... ... .     11
PL3107  /// /.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... A.. ... ... ... ... ..A C.. ... ... ... ... ... .      1
```

(b) HCV HVR1 DEDUCED AMINO ACID SEQUENCES: L3

```
          + 1                  10                  20                  No. of
          |                    |                   |                   clones
PL3001    G T Y T T G G A A G Q T T S G L A G L F N S G A R Q H      9
PL3022    . . . . . . . . . . . . . R . . . . . . T . . . . . .      1

PL3101    . . . . . . . . . . . . . R . . . . . . T . . . . . .     11
PL3107    / / . . . . . . . . . . . . . I . . . . K P . . . . .      1
```

Dots (.) indicate nucleotide (a) or deduced amino acid (b) sequence identical to dominant sequence of first time point. Silent mutations shown in lower case. Nt. 1 corresponds to nt 1489 of HC-J8 (Okamoto et al., 1992).

5.3 Long-term evolution of HVR1 quasispecies populations in immunosuppressed liver transplant recipients

In two patients (C1 and L2), HCV HVR1 sequences were determined from serial samples obtained over 5 time points spanning approximately one year. In both cases, the dominant molecular species found immediately post-OLT remained virtually unchanged over the observation period. The few detected mutants differed from the major sequence at no more than one of 27 deduced amino acid positions. This relative sequence homogeneity contrasted markedly with results from the asymptomatic HCV hosts Wad, Dix and Fra (patients 1 - 3), who had received no immunosuppressive treatment and were followed up over similar time intervals. In Wad (patient 1), after 9 months a previously-undetected variant accounted for 4 of 11 sequences (Figure 5.1). Similarly, in Dix (patient 2, see figure 5.2), after 14 months, the initially-dominant variant had become undetectable, while the minor variant had acquired a substantial number of mutations. In both untreated cases, these mutations effected nine deduced amino acid substitutions of the 27 total amino acids comprising the HVR1 (Figures 5.1(b) and 5.2(b)).

5.4 HVR1 quasispecies distribution pre- and immediately post-transplantation

The HVR1 sequences of C1, L3 and L1 (patients 4, 6 and 7, respectively) were determined within the month preceding and 2 - 3 months post-OLT. In each case, a minor species or closely related variant of the pre-OLT population became dominant or exclusive post-OLT. However, the total number of variants was reduced only in one patient (C1), from 6 pre-OLT to a single variant at time points one week and 3 months post-OLT (Table 5.2). The number of variants remained the same in L3, and increased from 3 to 6 in L1. In C1, only one of 5 observed mutations led to an amino acid replacement, which was a conservative Ile to Leu change (figure 5.5(b)). In both C1 and L3, mutations of pre-OLT minor variants produced deduced amino acid replacements that persisted post-OLT, which were still detectable in C1 after one year of follow-up. In L1 (patient 6), on the other hand, the dominant HVR1 sequence observed three months post-OLT (4 months after the pre-transplant sample) was

different from the pre-OLT minor species, albeit apparently derived from one of them (SLB015), with identical replacement mutations at four positions (figure 5.9).

Table 5.2: *HCV HVR1 molecular species in immunosuppressed and untreated subjects*

Subject	No. of quasispecies:					
	Pre-OLT	Post-transplantation				
		S1*	S2	S3	S4	S5
1 (Wad)	NA*	3	2			
2 (Dix)	NA	2	4	10		
3 (Fra)	NA	3	8	2		
4 (C1)	6	2	1	3	4	
5 (L2)	-*	5	3	2	2	4
6 (L1)	3	5				
7 (L3)	2	2				

NA = not applicable; (-) = not available; S1 etc. = sample #.

5.5 HVR1 mutation patterns in immunosuppressed and untreated subjects

The HVR1 nucleotide and amino acid sequence diversities of the 4 immunosuppressed and 3 untreated carriers of HCV were compared. As detailed above, the average number of genomic variants per sample was lower (range 1 - 6 , mean 3.14) in the immunosuppressed patients than in the untreated individuals (range 2 - 10, mean 4.25). The maximum nucleotide sequence diversity ranged from 1 - 19.8% for immunosuppressed patients (mean 2.3%), as compared with 2.5 - 30.9% (mean 20.2%) for untreated subjects. The contrast was more marked when considering sequence diversity at the putative amino acid level: Mean values were 2.3% in immunosuppressed patients and 43.5% in untreated HCV hosts. Tables of HVR1 nucleotide and amino acid sequence diversities in all possible pairwise combinations are given in appendix B. The values are for each set of clones, and include those between individual clones in different populations (from serial samples) obtained from the same individual. Table 5.3 gives the values for maximal nucleotide and amino acid sequence

diversities within the various populations of HCV HVR1 sequences obtained from the seven individuals discussed in this chapter. When the periods of observation were summed up for each group of patients, non-immunosuppressed individuals and OLT recipients had an average of 2.58 and 1.61 substitutions per month, respectively (Table 5.3). These mutations corresponded to 1.94 and 0.4 deduced amino acid substitutions/month in the untreated and immunosuppressed groups, respectively.

Table 5.3: *Maximum % nucleotide (amino acid) differences within populations of HCV HVR1 sequences from immunosuppressed patients and untreated HCV infected subjects*

	Pre-OLT	Post-transplantation				
Host	**S0***	**S1**	**S2**	**S3**	**S4**	**S5**
Immunocompetent:						
Wad	N.A.	2.5 (3.7)	17.3 (33.3)			
Dix	N.A.	13.6 (25.7)	22.2 (55.6)	23.5 (51.9)		
Fra	N.A.	23.5 (51.9)	30.9 (66.7)	16.27 (59.2)		
Immunosuppressed:						
C1	8.6 (3.7)	1.2 (0)	0 (0)	2.5 (3.7)	2.5 (0)	
L1	9.9 (22.2)	7.4 (14.8)				
L2	N.A.	19.8 (29.6)	4.9 (0)	1.2 (0)	1.2 (0)	3.7 (3.7)
L3	3.7 (7.4)*	4.9 (14.8)				

** S0 to S5 refer to individual sample HVR1 populations; N.A. = not applicable*

Other parameters of comparison between the two groups of HCV-infected individuals were investigated. As shown in Table 5.4, a significant difference was found between the ratios of replacement (R) and silent (S) mutations (R/S ratios), which were 8.22 in untreated subjects, versus 1.33 in transplant recipients ($p < 0.01$). The ratios of transitions (T_S) to transversions (T_V) were also significantly different, at 0.98 and 2.57 for untreated and treated subjects, respectively ($p < 0.02$). There was also a difference in ratios of conservative vs non-conservative amino acid substitutions between samples taken from immunosuppressed patients (61:34 or 1.79) and untreated individuals (15:10 or 1.5).

Table 5.4: *HCV HVR-1 mutations in untreated and immunosuppressed subjects*

	Untreated	Immunosuppressed
Total no. of nt mutations per month[1]	2.58	1.61
Transition/transversion ratio	1.21	2.57
Average no. of quasispecies	4.25	3.14
R:S ratio	3.70	1.21
Total no. of aa replacements	1.75	0.94
Conservative:non-conserv. aa replacement ratio[2]	1.79	1.50

[1] In each group, the time intervals between the first and last samples were summed up for all the patients.

[2] For the immunosuppressed group, pre-OLT (i.e., pre-immunosuppression) data were left out of the calculations.

When comparing pre-transplant samples from OLT recipients with the first available samples from untreated individuals, the number of molecular species was similar in both groups. The maximum number of nucleotide substitutions was higher in the untreated group (13.2 and 7.4 for untreated and immunosuppressed individuals, respectively). This difference, due to the large difference in the ratio of replacement to silent mutations (R/S ratio) between the two groups (9.7 vs 1.4 in untreated and transplanted subjects, respectively), became significant only at the amino acid level: 27.1% maximal amino acid sequence diversity in untreated subjects vs 11.2% in transplanted patients.

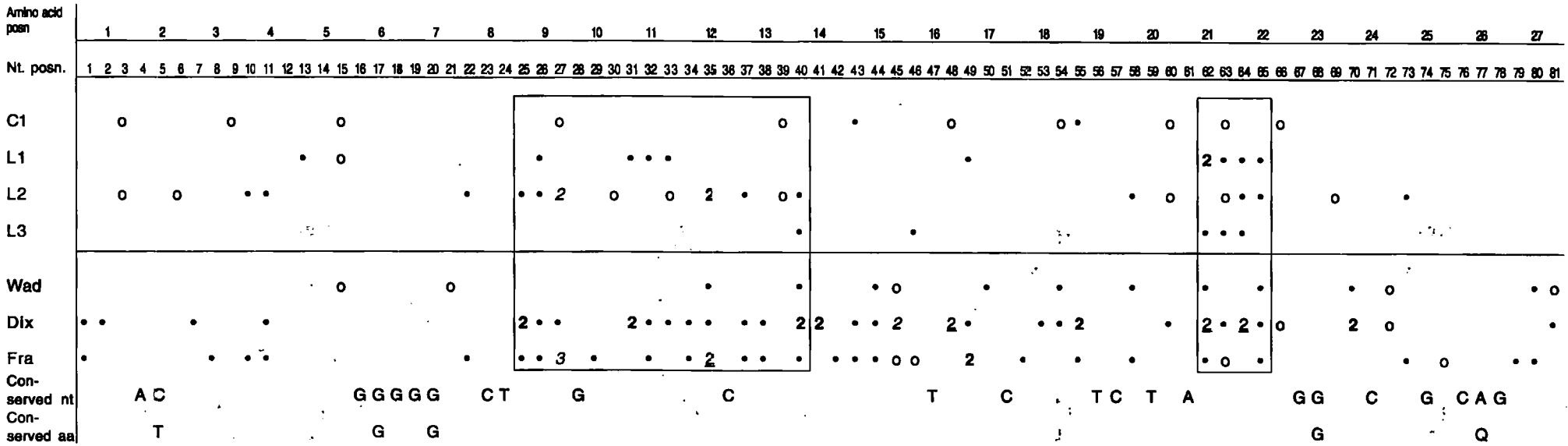
5.6 Localisation of HVR1 nucleotide mutations

The HVR1 sequences described in this chapter totalled 255 - 85 from untreated asymptomatic HCV-infected individuals and 170 from transplanted patients. Each mutated HVR1 position was scored, and the values depicted as in figure 5.11. From this, an interesting pattern of HVR1 mutability emerged. Of the 81 nucleotide positions of the HVR1 sequence, 25 (31%)

were absolutely conserved, 12 of them encoding 5 inferred amino acids at HVR1 positions 2, 6, 7, 23 and 26 (Thr, Gly, Gly, Gly and Gln, respectively).

Some degree of variability was observed in all other areas of the HVR1. In particular, analysis of sequences from both groups of patients revealed two mutational hotspots, located at HVR1 positions 25-40, and 62-65. The mutation rate was particularly high in the latter region, and led mostly to amino acid replacements. A difference between the two groups was found in the region spanning nucleotide positions 43-55: a marked drop in variability of this area occurred only in the immunosuppressed group. Similarly, the variability of nucleotides 70-81 in untreated patients, albeit lower than in the hotspots, far exceeded that found in immunosuppressed patients. At the first mutational hotspot, R/S ratios were 6.25 and 1.57 for immunocompetent and immunosuppressed patients, respectively ($p = 0.0024$). In contrast, there was no significant difference in R/S values between the two groups at the second hotspot. Multiple mutations at a single nucleotide position were found in two immunosuppressed and two immunocompetent patients. They were more common in the untreated group.

Figure 5.11: Schematic depiction of nucleotide substitutions in the HCV HVR derived from 7 infected patients who received (C1, L1; L2, L3) or did not receive(Wad, Dix, Fra) immunosuppressive treatment (patients 1 - 4, and 5 - 7, respectively).



Open and filled-in circles (o and •) denote, respectively, silent and replacement mutations. The figure 2 emboldened or italicised indicates 2 replacement or two silent mutations, respectively. When underlined, it depicts one silent and one replacement mutation.

DISCUSSION

CHAPTER SIX

This investigation revealed marked differences between the patterns of HCV mutability of immunosuppressed patients and non-transplanted individuals. Due to the large number of nucleotide sequencing reactions carried out in this investigation, care was taken to optimise individual steps in the protocol to obtain maximum efficiency. This included a study of the effect of various anticoagulants, and of stringent RNA denaturation on HCV sequence amplification from plasma, described below.

Differential effects of heparin on HCV NCR and E1/E2 amplification:-

A striking finding was the apparent selective inhibition of PCR by heparin. As discussed in section 4.1, amplification of the HCV *E1/E2* region was not possible when plasma derived from heparinised blood was used. However, heparin appeared to have no inhibitory effect on amplification of NCR sequences. In seven consecutive isolates from one patient, for example, the HCV NCR was repeatedly amplified, while *E1/E2* PCR was negative. Similarly, the use of heparinised blood prevented amplification of *E1/E2*, but not NCR, sequences from samples obtained from other patients. That this effect was differential, and that it was due to heparin, was implied by the fact that non-heparinised samples obtained from time points close to those for heparinised samples were consistently PCR-positive for *E1/E2*.

Heparin is a proteoglycan composed of a glycosaminoglycan (GAG) chain which is covalently linked to a single-chain protein core. The mechanism whereby it inhibits the various steps of RT- PCR is unclear. It is highly negatively charged due to its many sulphate and carboxyl groups, suggesting that it may compete with DNA (or RNA), also negatively-charged, for binding to the enzymes reverse transcriptase or *Taq* polymerase. Alternatively, it may bind to nucleic acid, preventing polymerase activity. Binding may be mediated by cations or by positively-charged portions of the component polypeptide core. There is evidence suggesting that the inhibitory effect of heparin is more pronounced in samples exposed to heparin for

longer periods (Satsangi *et al.* 1994), implying that DNA binding is the primary mechanism whereby it exerts its effects.

The differential inhibitory effect of heparin observed in this study may be due to topological constraints on binding of heparin to the NCR, which is believed to have a complex secondary, and perhaps tertiary, structure including a configuration characteristic of an internal ribosome entry site. However, it is impossible to state definitively that this is the case in the absence of quantitative data on viraemia levels in the samples analysed. Factors affecting PCR outcome include the size and location of the targeted amplicon. The size of the NCR amplicon in these experiments was less than half that of the *E1/E2* amplicon. Furthermore, the NCR is located at the end of the RNA molecule, which would render it more easily accessible to the enzymes involved in RT-PCR. Hence the observed positive results for NCR amplification may reflect the fraction of total template NCR segments that escaped binding by heparin, rather than a lack of inhibitory effect in this region, and quantitative data may show this effect.

In a single heparinised sample (C1), it was possible to obtain an *E1/E2* PCR band from RNA obtained using the magnetic bead method described in section 2.7, but not from RNA obtained by guanidinium thiocyanate denaturation followed by alcohol precipitation. The positive *E1/E2* PCR result from a heparinised sample indicated that the magnetic bead method, which involves selective binding of HCV RNA to, and subsequent elution from, streptavidin-coated magnetic microparticles, is able to reduce or eliminate the inhibitory effect of heparin on RT/PCR. It is likely that extensive washing of the beads pre-elution of RNA dislodged the heparin from the RNA. The effect was not reproducible, and for further studies it was necessary to use non-heparinised samples.

Importance of stringent RNA denaturation conditions for E1/E2 amplification:-

Another important factor in successful amplification of HCV cDNA was the application of stringent RNA denaturation conditions, particularly for *E1/E2* PCR. Although single-

stranded, RNA contains local regions of short complementary base pairing that can form from a random matching process. Standard protocols for RT-PCR recommend a 5 min, 65°C incubation of the RNA before reverse transcription (Innis *et al.* 1990). RNA pre-incubation regimes described in studies involving amplification of the HCV non-coding region range from room temperature for 15 min (Bouffard *et al.* 1992) to 94°C for 5 min (Hu *et al.* 1993). In reports describing amplification of HCV *E1/E2* sequences, the corresponding regimes were 65°C for 5 min (Feray *et al.* 1992); 70°C for 5 min (Ogata *et al.* 1991; Driesel *et al.* 1994); and 80°C for 5 min (Van Doom *et al.* 1994). Following the fortuitous discovery in this study that stringent heating of HCV RNA before cDNA synthesis strongly influenced the outcome of RT-PCR, a series of experiments, described in section 3, was carried out to validate this observation. Different heat and chemical denaturation protocols were compared. The finding that stringent RNA heat or chemical denaturation conditions (70°C for at least 30 min or 35 s boiling of formamide-dissolved RNA) were necessary for positive *E1/E2* PCR suggests a significant degree of higher order structural organisation for this region of HCV.

Genotyping of HCV via amino acid sequence homology in the E1 carboxyterminus:-

The phylogenetic classification system of Simmonds *et al.* (1995), involving "types" and "subtypes", was generated by computer. While isolates within subtypes had only 0 - 12% nucleotide variation, subtypes within types had 14 - 26% difference. Each of the major genotypes differed by 28 - 44%. The three-tiered system was supported by the fact that analysis of several subgenomic regions revealed the same subdivisions, with no overlap in sequence variability between isolates, subtypes and types. The data suggested that genetic recombination between different HCV genotypes does not occur, and that classification can reliably be performed by analysis of informative subgenomic regions. This was the basis of classification in the Bukh-Okamoto system (Bukh *et al.* 1993; Okamoto *et al.* 1992), used in this study. The region for distinguishing between HCV genotypes in most cases could be narrowed down to a 10-amino acid domain from residues 342 - 351, which was previously described as sufficient to reliably type a large number of samples obtained from sources worldwide (Bukh *et al.* 1993).

In the genotype determinations of chapter 4, amino acids 342 to 351 were able to type 12 of 12 samples. An apparent ambiguity with one sample (C1) was resolved by the finding, after comparing the 7 subtype 1a and 6 subtype 1b sequences of figures 4.5 and 4.6 respectively, as well as the 8 subtype 1a and 17 subtype 1a sequences of Bukh *et al.,* that subtype 1a could be distinguished by the invariable presence of Leu at position 345. This was confirmed by analysis of the regions flanking this 10-residue sequence, in which C1 was more homologous with the subtype 1b. Because the region of *E1* used for typing was located close to the HVR1, it was possible to classify HCV isolates and carry out HVR1 analysis on the same PCR product, which significantly increased efficiency.

Independence of HVR1 variability from genotypes: Implications for cross-reactivity and neutralization:-

The results of section 4.1 indicated an apparent absence of subtype specificity in the HVR1 region. However, constraints on the identity and hydropathic character of amino acids comprising HVR1 were found. These included three absolutely conserved residues: Thr, Gly and Gln at HVR1 positions 2, 23 and 26, respectively. In addition, in 92% of isolates, Gly was found at both positions 6 and 7, Phe at position 21 and the hydrophobic residues Leu or Phe at position 20. These results suggest a degree of conservation within the HVR1 region which is independent of HCV genotype, and point to a possible functional role for this region of *E2*, such as binding to a cellular receptor. Although the relatively high degree of conservation expected from a cellular receptor binding site is absent from the HVR1 as a whole, conserved residues within it may interact with parts of E2 outside the HVR1 to constitute the functional binding site. In this context, it was recently demonstrated that anti-HVR1 antibodies were able to block recombinant *E2* binding to susceptible MOLT-4 cells (Rosa *et al* 1996). The presence of conserved residues within the HVR1 may have implications in the design of genetically engineered neutralising antibodies to HCV, which would widely cross-react with multiple subtypes. Progress in passive immunity as well as vaccine development may benefit from further determinations of consensus sequences from a large number of isolates, and

clarification of the structural role of the conserved amino acids (which may specify widely-conserved tertiary structural motifs).

Observation of two intraisolate viral "families" drifting in parallel:-

The relatively frequent observation, among both transplanted and untreated subjects in this study, of two viral "families" apparently drifting in parallel was of note. This phenomenon could be related to initial infection with two equally infectious variants; superinfection over time; or adaptation to host. The simultaneous presence of HCV from different genotypes is not uncommon. However, experiments performed in chimpanzees successively infected with different strains of viruses of the same genotype suggest that, over time, one strain seems to take over (Farci *et al.* 1991). Hence, the hypothesis of a dual infection is unlikely. Superinfection of a previously infected individual by another strain of HCV has been observed in humans (Prince *et al.* 1992; Kao *et al.* 1993). Experiments in chimpanzees clearly showed that the level and the efficacy of neutralizing antibodies against HCV was insufficient to prevent superinfection (Farci *et al.* 1992; Okamoto *et al.* 1994). Untreated subject Wad (patient 1), in whom parallel drifting of two viral families was observed, had been infected through intravenous drug abuse more than 15 years before the study, and recent superinfection is very unlikely. There remains the possibility of two strains of escape mutants having drifted separately over a relatively long period of time. It is not inconceivable that two distinct mutated strains have become well adapted to the host and persisted.

Decrease in HCV genomic diversity post-OLT:-

The data collected in this study indicate a considerable reduction in HVR1 genomic diversity of HCV isolated from liver transplant recipients when compared with isolates from chronically infected, untreated patients. This was especially remarkable in two of four transplanted patients (C1 and L2, or patients 4 and 5, respectively). The HVR1 quasispecies populations in both patients had become practically homogenous by one year post-transplantation. As shown in Figures 5.1 - 5.9 and Tables 5.1 and 5.2, the overall number of variants, as well as the frequency, type and location of nucleotide mutations appeared

restricted post-transplantation. These observations can be explained by the following mechanisms:

- restriction of infectious virus species infecting the grafted liver;
- primary infection of grafted liver;
- convergent evolution;
- attenuated host immune response;
- low viral diversity preceding liver transplantation.

Each of these possible factors is discussed below:

Restriction of infectious viral species:- The basic premise here is that only one or a few variants is capable of infecting the grafted liver. In 2 patients (C1 (patient 4) and L3 (patient 7)), a minor pre-OLT HVR1 species became unique or dominant post-OLT. Similarly, for patient 5 (L2), a minor species observed 31 days post-OLT dominated all subsequent samples collected over a year. This apparent selection may be explained by a restriction in the number of viral species capable of, or available for, infecting the transplanted liver. Preliminary evidence has been provided that HCV in plasma circulates in two forms - as free virions or as virus-antibody immune complexes (Choo *et al.* 1995). Evidence from other investigators suggests that complexed viruses have decreased or no infectivity, while free viruses probably correspond to escape mutants not recognised by the immune system (Shimizu *et al.* 1994; Zibert *et al.* 1995). This hypothesis is supported by the observation that HVR1 sequences of complexed and free HCV in plasma have different nucleotide and amino acid compositions. In liver graft recipients, free viruses may represent a selected, more infectious viral population which preferentially infects the grafted organ. During this primary infection, a limited number of HCV variants replicate, producing viruses of limited diversity. This argument is based on the assumption that the viruses derived from plasma or serum represent an accurate picture of the hepatic HCV population. In the absence of good tissue culture systems for HCV, it is difficult to determine which genomic characteristics determine infectivity. Removing the infected liver of potential graft recipients would remove most of the virus. Selection may then operate

on the basis of those variants surviving outside the liver (such as in PBMCs, or virions in circulation at the time of the operation).

Primary infection of grafted liver:- There is evidence to suggest that a relative homogeneity of viral genotype in the period immediately following virus transmission is a fairly common phenomenon (Feray *et al.* 1992; Gretch *et al.* 1995). Thus, a single predominant HVR1 variant was found in the infant of a mother shown to have multiple predominant HVR1 species (Weiner *et al.* 1993). The infant sequence was closely related to, but not identical with, the 9 variants identified in the mother at the time of birth, a situation closely mirroring that found in OLT recipient C1 (patient 4) in this study (figure 5.5). There appears to be a "bottleneck" effect in the transmission of virus so that the preferential replication of one or a few variants is favoured. The particular species transmitted could result from a completely random event, or may have a better replicative ability. In the latter context, it is notable that a dramatic increase in HCV viraemia in the period following OLT of previously infected patients has been reported from several investigators (Alter and Seeff 1993; Duvoux *et al.* 1994; Gretch *et al.* 1995). For a short period of time (1 - 2 months), during which restricted infectious HCV species explosively replicate in the target organ, a very high proportion of virus particles produced are identical to the species with replicative advantage. Competition experiments between mutant and wild-type Qβ RNA phages showed that mutants with a relative replication rate of up to 0.9 were nevertheless outgrown by the wild type, which formed the predominant sequence (Holland *et al.* 1982). Ordinary sequencing methods would detect RNA heterogeneity only if the mutant-wild type growth ratios were very close (within 0.01 - 0.1%). This mechanism, applied to the post-OLT context, assumes that viral diversity would begin to reappear 4 - 6 months post-OLT (1 - 2 months following primary infection of the grafted liver), when the log phase of growth is over, and dynamic equilibrium of quasispecies variants may have shifted to favour the predominance of one or more new species. This is not the case in C1 and L2 (patients 4 and 5), who still show restriction of viral genomic diversity 1 year post-OLT, despite evidence, from quantitation studies, of increased viral load (see appendix D).

Convergent evolution:- The final antigenic character of a given viral epitope results from a balance between two opposing trends, whereby the variability which enables escape from the host immune response is weighted against the need to conserve specific residues that are important for viral function. Hence, two or more distinct evolutionary lineages may give rise to a single antigenic epitope (Holmes *et al.* 1992). The apparent homogeneity of HCV variants in patients C1 and L2 may thus represent convergent evolution of different lineages to produce a preponderance of that variant which is best fit for survival within the specific host.

Decrease in the host immunological ability to select HCV strains:- In immunocompetent individuals, the emergence of new HCV variants appears to be essentially the result of immunologic selection. Dominant variants are quickly recognised by the immune system, which produces antibodies against antigenic motifs specific to each variant (anti-envelope, particularly anti-HVR1). Clearance of these complexed viruses allows the emergence of variants which are not immunologically recognised (escape mutants), and which circulate at levels high enough to be detected amongst the limited number of sequenced clones. The immunosuppressive regimen undergone by OLT recipients may interfere with their ability to recognise escape mutants and to mount an effective immune response. If this is the case, patients will have a decreased or non-existent capacity to select HCV variants as escape mutants. Since the immunoselective mechanism is impaired, the initially infectious species is (are) likely to keep replicating and to remain longer in circulation than prior to the onset of immunosuppression. Thus, only after a year post-OLT did the HCV population from patient 4 (C1) begin to diversify, with the few observed mutations mostly silent.

Pre-existing relative inability of transplanted patients to select divergent variants:- In three patients examined pre-OLT several indicators suggest that genomic diversity is less than in the three untreated patients. In C1, for example, although 6 species were detected pre-OLT, only two variants were represented at the amino acid level. Similarly, for the remaining transplant recipients, the degree of amino acid diversity was lower than that

found in the non-transplanted subjects. The maximum percentage nucleotide diversity in the immunosuppressed patients was 9.9%, compared with 23.5% in untreated subjects (corresponding to maximal amino acid sequence divergences of 22.2% and 52%, respectively). The ratios of transitional to transversional mutations (Ts/Tv ratios) were also higher in immunosuppressed patients. This scenario differs from that in the preceding paragraph in that here the restriction on immune selection of divergent strains is intrinsic to the individual patient (e.g., to a specific HLA haplotype), and not a result of the imposed immunosuppression. These observations suggest that patients who develop end-stage liver disease have a lower HCV genomic diversity than those who do not, which might reflect their decreased capacity of their immune systems to recognise viral antigenic variants and select quasispecies populations. More data are necessary to substantiate this point.

Two mechanisms of virus evolution have been proposed. Random mutations occur in the absence of evolutionary pressure. In this process, the main viral species remains essentially unchanged, with a gradual drift over long periods of time. Alternatively, in the presence of selective pressure, such as a host immune response, mutants which differ antigenically from the major variant and have not yet attracted the attention of the host immune response are selected. New, sometimes drastically different, species sequentially emerge as escape mutants which may ultimately be identified and eliminated by the host immune system.

Evidence for both kinds of mechanism was found in this study, as outlined in chapter 5. The HVR1 sequences of patients 4 and 5 (C1 and L2) remained remarkably stable throughout the one-year follow-up period (figures 5.5 and 5.7). Selective forces appeared to operate in the untreated subjects Wad, Dix and Fra (patients 1 - 3, figures 5.1 to 5.3, respectively), as evidenced by the rapid turnover of significantly different quasispecies populations. The mechanism of evolution in the two immunosuppressed OLT recipients L1 (patient 6) and L3 (patient 7) was difficult to determine, as a long-term study was not carried out on these patients. However, from the sequencing data for the initial post-OLT sample from each

patient, selective pressure appeared to play some role. As seen from figures 5.9 and 5.10, there were significant differences in antigenic potential among the post-OLT HVR1 genomic variants in both patients. Hence, the possibility of immune selection cannot be ruled out for all OLT patients.

The HVR1 is by far the most variable part of the HCV genome, with a very high degree of sequence heterogeneity detectable among multiple clones (Ogata *et al.* 1991; Weiner *et al.* 1991; Honda *et al.* 1994; Kato *et al.* 1994). Variant composition was found to change temporally, in a process involving selection of particular amino acids at specific positions which had represented only a minor fraction or had been absent in previous clones. Furthermore, this heterogeneity could be decreased by IFN treatment (Higashi *et al.* 1993). RNA-dept RNA polymerases have a bias towards base transitions, implying that HVR1 is not a section of redundant RNA that arose due to random infidelity of the virus polymerase, with no constrainst on its sequence. The segmental evolution of HCV must have undergone positive selection. The low Ts/Tv ratio in HVR1 highlights the possibility that the mechanism for generation of this diversity involved a different replication strategy. The finding of constant regions and hotspots interspersed within HVR1 further highlights the critical role of this region in the immune response.

The HCV subtype 1b is believed to be more pathogenic than other subtypes (Pozzato *et al.* 1994). However, this could merely reflect the length of incubation rather than viral pathogenicity. HCV infection follows a slow, insidious course, with evidence of liver disease often taking decades to develop. The lack of extensive long-term studies could mask the possibility that the subtype 1b (which, incidentally, is widely distributed globally) has simply been present much longer. In this context, it may be significant that the population of patients in this study infected with subtype 1b was, on average, 20 years older than those with other subtypes.

In two untreated subjects, Wad and Dix, 14 and 23 HVR1 mutations, respectively, were observed between sequences from two time points. This was consistent with results from a 12-patient study in which 7 - 15 HVR1 mutations (median 10) were observed in consensus sequences obtained from paired samples separated by a one-year interval (Kurosaki *et al.* 1994). The numbers of mutations observed in immunosuppressed patients C1 and L2 were lower (6 and 4, respectively). Temporal changes in deduced amino acids in the untreated subjects also corroborated other published reports, and were clearly higher than changes observed in the immunosuppressed OLT recipients (Higashi *et al.* 1993; Kato *et al.* 1994 ; Kao *et al.* 1995) .

Comparison of the two groups of patients was even more striking when samples collected at relatively close intervals over periods of several months were analysed. In untreated subject Fra for example, amino acid sequence divergence reached 67% in less than a month, while two immunosuppressed patients 4 and 5 (C1 and L2) showed a maximum change of 4% throughout a follow-up period of almost one year. Two overlapping epitopes, corresponding to amino acid positions 11 - 21 (epitope I) and 14 - 24 (epitope II) of the HVR1, have been previously described (Kato *et al.* 1994). It is of interest to note a general correspondence between these epitopes and the mutational hotspots found in both groups of patients in this study. It is also striking that the decrease in mutability in immunosuppressed patients was most conspicuous in the overlap region of epitopes I and II.

All criteria examined for the comparison of HVR1 sequences revealed a difference between immunosuppressed and immunocompetent patients. The average number of variants, nucleotide mutation rate and sequence diversity were lower in the immunosuppressed group. In addition, mutations in immunosuppressed patients tended to be mostly transitional and more often silent, while replacement mutations tended to favour conservative substitutes. These elements suggest that a normal immune system plays a critical role in the determination of the quasispecies distribution of HCV infected individuals. It is unlikely, however, that immunosuppression alone accounts for this remarkable difference. One possible additional

influence is the individual patient's immune repertoire - patients who reach the stage of terminal liver disease at a relatively young age may represent a selection of people whose immune response, both humoral and cellular, is particularly inefficient. Limited data obtained from different chimpanzees inoculated with the same viral population suggest that an unfavourable clinical outcome is correlated with low levels of HVR1 variability (Van Doorn *et al.* 1994; Van Doorn *et al.* 1995). The low HVR1 variability observed in the OLT recipients described in his study may have preceded the period of immunosuppression and may be related to the end stage liver disease requiring liver transplantation. It may have been caused by the host inability to contain HCV (hence the cirrhotic state), or to the immunosuppressive treatment given for transplantation, or both.

These data need to be substantiated in more patients, both OLT recipients and untreated, chronically infected subjects. However, they suggest that patients who develop end-stage cirrhosis have a lower genomic diversity of HCV, which might reflect a decreased capacity of the host immune system to recognise viral antigenic diversity and to complex HCV. Patients 4, 5 and 7 who developed end-stage cirrhosis at a relatively young age, may represent a subgroup of HCV-infected individuals whose immunological response is ineffective, hence the lower viral diversity and genetic drift. In addition, it should be noted that C1 (patient 4), who had the lowest level of genomic diversity, rapidly developed HCV-related graft failure, requiring re-transplantation after one year. In a Scandinavian study, HCV-infected patients with congenital hypogammaglobulinaemia were shown to undergo a severe, rapidly progressive course of hepatitis, with poor response to IFN (Bjoro *et al.* 1994), which accords well with the findings for patients 4 (C1) and 5 (L2). Kumar *et al.* showed absolute homology of the HVR1 nucleotide sequence in an agammaglobulinaemic patient followed for two years (Kumar *et al.* 1994). These results indicated that HCV persistence occurs in the presence or absence of a humoral immune response, and suggested a role for other factors in promoting viral persistence, such as the cellular immune response. Evidence from several sources points to the existence of T cell responses directed against HIV, a virus which, like HCV, is remarkable for its high rate of mutation in the region encoding envelope

glycoproteins. (McMichael *et al.* 1994; Klenerman *et al.* 1995; McAdam *et al.* 1995; RowlandJones *et al.* 1995). Studies on HCV have produced similar findings (Kita *et al.* 1993; Koziel *et al.* 1993; Shirai *et al.* 1994; Shirai *et al.* 1995).

Although the number of patients in this investigation was not high enough to draw statistically relevant conclusions, it was notable that the two transplanted patients followed up for a year, who were both infected with HCV subtype 1b and showed a similar pattern of HCV quasispecies evolution, had identical haplotypes at the HLA-A locus. The fact that they both required re-transplantation within a year of the first transplant further suggests a possible link between host factors and the increasingly evident subset of HCV-infected OLT recipients who develop very aggressive liver disease post-transplantation.

The data presented suggest a model for the difference in HCV HVR1 evolution between untreated and subjects at least a subset of immunosuppressed patients (figures 6.1 and 6.2). At any time point, the viral population is composed of multiple variants. In the kinetics of these molecular species, only mutations affecting epitopes and, particularly, functional epitopes of the viral structural proteins are of importance in the host-virus interaction mediated by the immune system. This is especially relevant for the HVR1 region. Viral particles circulate either as free, infectious virions or complexed with specific antibodies, which prevent their replication within cells and are therefore neutralising. Variants generated during the replication process are eliminated in immune complexes, or temporarily evade neutralisation. The constant clearance from circulation of immunogenic variants complexed with specific antibodies allows new variants to emerge and replicate for various periods of time in sufficient amounts to become detectable as quasispecies variants.

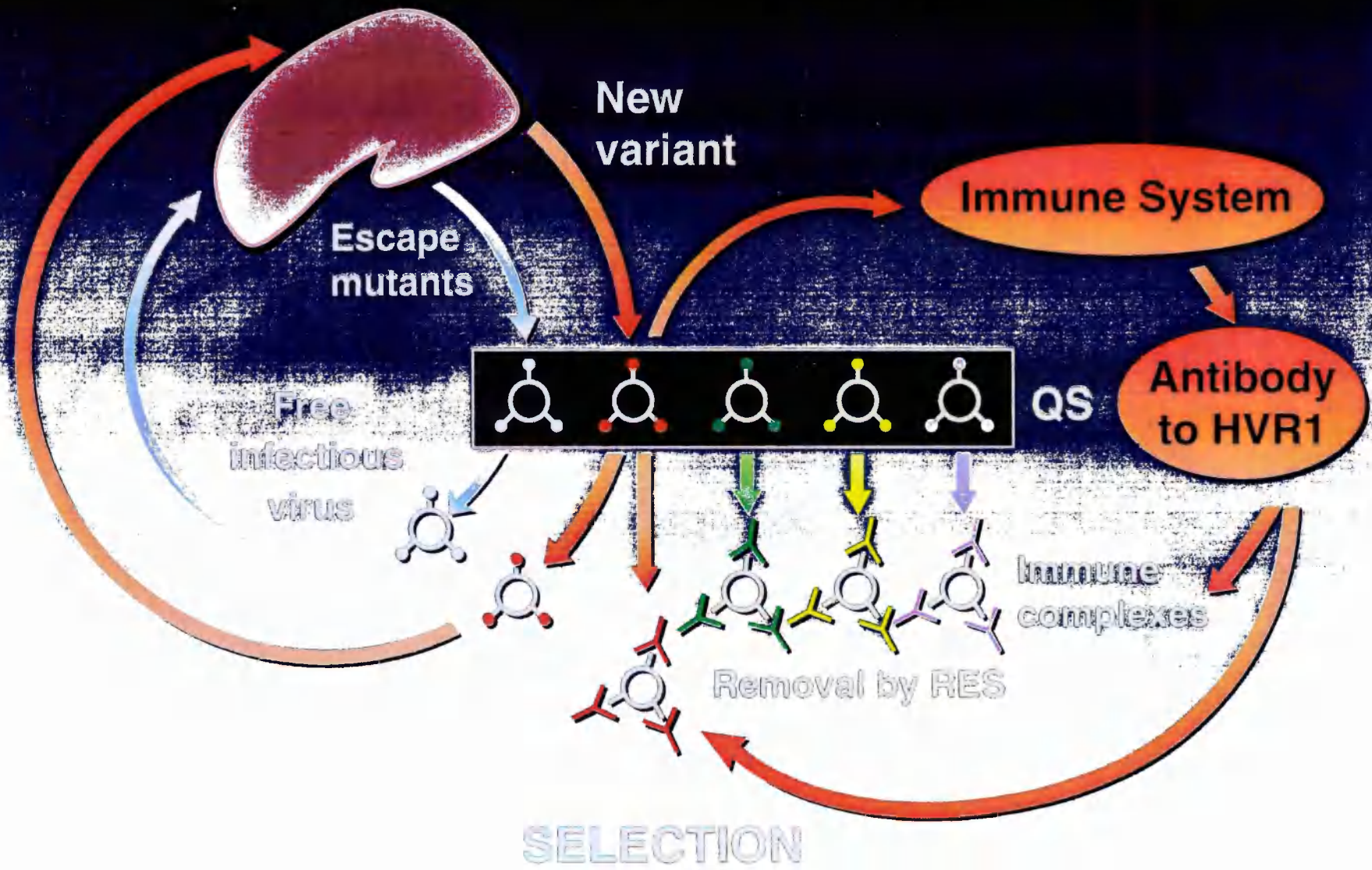
The model predicts that in the context of immunosuppression or a failure of the immune system to recognise and clear HCV variants, the most infectious mutant (i.e., uncomplexed escape mutant) will preferentially replicate, leading to a homogeneous viral population. Under these conditions, although other species may be present, their lack of infectivitiy advantage

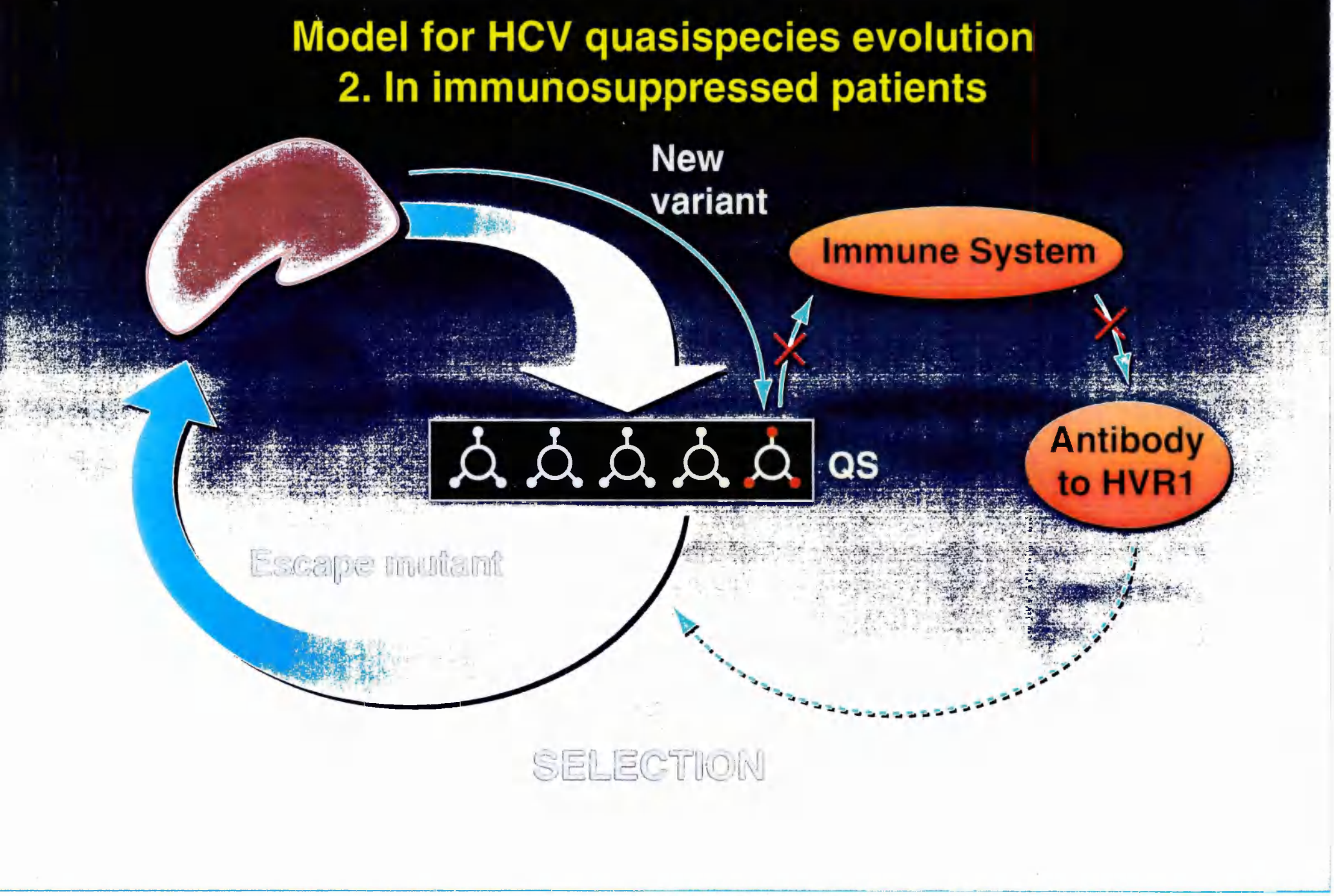
over the escape mutant which colonises the allograft means that they are present at levels too low for detection by the methods used in this study.

It should be emphasized that this model represents but one of the many possible mechanisms of viral quasispecies evolution, the one which best fits the situation found in patients C1 and L2. These patients were followed up long-term, unlike patients L1 and L3, for whom a single post-OLT sample was analyzed. For the latter two patients, these preliminary findings suggest that HVR1 sequence heterogeneity, high in both cases, is predictive of an uncomplicated post-OLT course. While the two patients with relatively high HCV HVR1 sequence homogeneity remain well over two years post-OLT, both patients with homogenous quasispecies populations developed severe recurrent hepatitis which proved lethal despite re-transplantation carried out within a year of the first transplant, .

Extensive clinical data review is required to substantiate the hypothesis of differential quasispecies evolution patterns in different subsets of liver transplant recipients, especially in view of its possible high impact on OLT outcome. Although IFN therapy, currently the only treatment for HCV infection, is expensive and of low effectiveness, the application of combined IFN/ribavirin therapy is reported to have significantly better results, even within the context of OLT and its attendant complications. As more effective therapies for HCV infection emerge, it will become increasingly important to pinpoint those patients most likely to suffer a complicated post-OLT clinical course, in order to assist selection of OLT recipients and/or choice of patients in whom more aggressive anti-HCV therapy should be carried out.

Model for HCV quasispecies evolution
1. In immunocompetent patients




Model for HCV quasispecies evolution
2. In immunosuppressed patients
New variant
Immune System
Antibody to HVR1
QS
Escape mutant
SELECTION

CONCLUSION

In conclusion, evidence is presented for different patterns of HCV sequence evolution in untreated subjects and some immunosuppressed patients. On the basis of HVR1 sequences, it is suggested that viral diversity is largely driven by immune pressure. However, the data indicate that patients showing less diversity of HVR1 sequences post-OLT may progress to a more severe clinical disease. The data validate and substantiate findings that the occurrence of changes in the HVR1, which encodes epitopes for neutralising bodies, reflect the efficacy of the immune system. Longer follow-up studies involving larger cohorts of patients are needed to clarify which viral and host factors are predictive of clinical outcome of infection in both asymptomatic HCV carriers and HCV-infected OLT candidates. In particular, clarification of the role of the cellular immune response, and the host immune repertoire, may help in the design of therapeutic strategies for the prevention and possible cure of hepatitis C.

REFERENCES

1. Agnello, V. (1995). "Mixed cryoglobulinemia and hepatitis C virus." Hospital Practice **30**(3): 35-42.

2. Agnello, V., R. Chung and L. Kaplan (1992). "A role for hepatitis C virus infection in Type II cryoglobulinemia." New England Journal of Medicine **327**: 1490-1495.

3. Alberti, A. and L. Chemello (1995). IFN naive and the relationship of HCV genotypes: response with combination therapy. EASL 30th conference, Copenhagen 19-23 August.

4. Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts and J. Watson (1983). Molecular Biology of the Cell. New York, Garland Publishing, Inc.

5. Allain, J., S. Dailey, Y. Laurian, D. Vallari, A. Rafowicz, S. Desai and S. Devare (1991). "Evidence for persistent hepatitis C virus (HCV) infection in hemophiliacs." Journal of Clinical Investigation **88**: 1672-1679.

6. Alter, H. and L. Seeff (1993). 29. Transfusion-associated hepatitis. Viral Hepatitis: Scientific Basis & Clinical Management. Edinburgh, Churchill Livingstone. 467-499.

7. Alter, M., P. Coleman, W. Alexander, E. Kramer, J. Miller, E. Mandel, S. Hadler and H. Margolis (1989). "Importance of heterosexual activity in the transmission of hepatitis B and non-A, non-B hepatitis." Journal of the American Medical Association **262**: 1201-1205.

8. Alter, H. (1988). Transfusion-associated non-A, non-B hepatitis: The first decade. Viral Hepatitis and Liver Disease. USA, Alan R. Liss, Inc. 537-542.

9. Arthur, M. (1994). "Cellular and molecular pathogenesis of liver fibrosis." Interaction **2**(3): 9-12.

10. Baltimore, D. (1971). "Viral genetic systems." Transactions of the New York Academy of Science **33**: 327-332

11. Barrera, J., M. Bruguera, M. Ercilla, C. Gil, R. Celis, M. Gil, M. Onorato, J. Rodes and A. Ordinas (1995). "Persistent hepatitis C viremia after acute self-limiting posttransfusion hepatitis C." Hepatology **21**: 639-644.

12. Bartolome, J., A. Madejon and V. Carreno (1995). "Ribozymes: Structure, characteristics and use as potential antiviral agents." Journal of Hepatology, Supplement **22**: 57-64.

13. BegicJaneva, A. and I. Boricic (1995). "HRONICNI HEPATITIS: STEPENOVANJE AKTIVNOSTI I ODREDIVANJE STADIJUMA. Chronic hepatitis: Grading of activity and staging." Archives of Gastroenterohepatology **14**: 148-153.

14. Belabes, E.-H., A. Bougermouh and J. Pillot (1988). Waterborne non-A, non-B hepatitis in Algeria: Epidemiological study and development of a test. Viral Hepatitis and Liver Disease. USA, Alan R. Liss, Inc. 152-153.

15. Belli, L., A. Alberti, G. Rondinara, L. De Carlis, F. Romani, G. Ideo and L. Belli (1993). "Recurrent HCV hepatitis after liver transplantation (24)." Lancet **341**: 378-379.

16. Bertoletti, A., A. Sette, F. Chisari, A. Penna, M. Levrero, M. De Carli, F. Fiaccadori and C. Ferrari (1994). "Natural variants of cytotoxic epitopes are T-cell receptor antagonists for antiviral cytotoxic T cells." Nature **369**: 407-410.

17. Bizollon, T., E. Sciarrino, C. Ducerf, U. Palazzo, M. Chevallier, M. Trabaud, J. Baulieux, M. Pouyet and C. Trepo (1995). Control of HCV recurrence following liver transplantation. EASL 30th conference, Copenhagen 19-23 August.

18. Bjoro, K., S. Froland, Z. Yung, H. Samdal and T. Haaland (1994). "Hepatitis C infection in patients with primary hypogammaglobulinemia after treatment with contaminated immune globulin." New England Journal of Medicine **331**: 1607-1611.

19. Bloch, K. (1992). "Cryoglobulinemia and hepatitis C virus." New England Journal of Medicine **327**: 1521-1522.

20. Blum, H. (1995). "Variants of hepatitis B, C and D viruses: Molecular biology and clinical significance." Digestion **56**: 85-95

21. Blumberg, B., B. Gerstley, D. Hungerford, W. London and A. Sutnick (1967). "A serum antigen (Australia antigen) in Down's syndrome, leukemia, and hepatitis." Annals of Internal Medicine **66**: 924-931.

22. Boker, K., G. Dalley, M. Bahr, H. Tillmann, C. Trautwein, H.-J. Maschek, R. Pilchlmayr and M. Manns (1995). "Long term outcome of liver transplantation for HCV-related liver disease." Journal of Hepatology **23**(Suppl 1): 153.

23. Boner, W., H. Schlicht, K. Hanrieder, E. Holmes and W. Carman (1995). "Further characterization of 2 types of precore variant hepatitis B virus isolates from Hong Kong." Journal of Infectious Diseases **171**: 1461-1467.

24. Bouffard, P., P. Hayashi, R. Acevedo, N. Levy and J. Zeldis (1992). "Hepatitis C virus is detected in a monocyte/macrophage subpopulation of peripheral blood mononuclear cells of infected patients." Journal of Infectious Diseases **166**: 1276-1280.

25. Brackmann, S., A. Gerritzen, J. Oldenburg, H. Brackmann and K. Schneweis (1993). "Search for intrafamilial transmission of hepatitis C virus in hemophilia patients." Blood **81**: 1077-1082.

26. Bradley, D. (1995). "Hepatitis E virus: A brief review of the biology, molecular virology, and immunology of a novel virus." Journal of Hepatology, Supplement **22**(1): 140-145.

27. Bradley, D., K. Krawczynski and M. Purdy (1993). 24. Hepatitis E virus: Epidemiology, natural history and experimental models. Viral Hepatitis: Scientific Basis & Clinical Management. Edinburgh, Churchill Livingstone. 379-383.

28. Bradley, D., K. Krzysztof, E. Cook, K. McCaustland, B. Robertson, C. Humphrey, M. Kane, J. Spelbring, I. Weisfuse, A. Andjaparidze, M: Balayan, H. Stetler and O. Velazquez (1988). Enterically-transmitted non-A, non-B hepatitis: Etiology of disease and laboratory studies in nonhuman primates. Viral Hepatitis and Liver Disease. USA, Alan R. Liss, Inc. 138-147.

29. Bradley, D. (1985). "The agents of non-A, non-B viral hepatitis." Journal of Virological Methods **10**: 307-319.

30. Brillanti, S. (1995). IFN relapsers and nonresponders: Italian experience. EASL 30th conference, Copenhagen 19-23 August.

31. Brillanti, S., M. Foli, S. Gaiani, C. Masci, M. Miglioli and L. Barbara (1993). "Persistent hepatitis C viraemia without liver disease." Lancet **341**: 464-465.

32. Brooks, S., E. Taylor, M. Golden and B. Golden (1991). "Electron microscopy of herpes simplex hepatitis with hepatocyte pulmonary embolization in kwashiorkor." Archives of Pathology and Laboratory Medicine **115**: 1247-1249.

33. Brown, E., H. Zhang, L. Ping and S. Lemon (1992). "Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs." Nucleic Acids Research **20**: 5041-5045.

34. Bruno, S., E. Silini, S. Rossi, F. Bono, F. Borzio, G. Leandro, M. Asti, A. Cerino, M. Pedicino, M. Mondelli and M. Podda (1995). "HCV genotype 1b is a major risk factor for the development of hepatocellular carcinoma (HCC) in cirrhotic patients: a prospective study." Journal of Hepatology **23**(Suppl 1): 111.

35. Bukh, J., R. Miller and R. Purcell (1995). "Defining the genotypes of hepatitis C virus and their epidemiology." INTER-action **3**: 10-15.

36. Bukh, J., R. Purcell and R. Miller (1992). "Importance of primer selection for the detection of hepatitis C virus RNA with the polymerase chain reaction assay." Proceedings of the National Academy of Sciences of the United States **89**: 187-191.

37. Bukh, J., R. Purcell and R. Miller (1993). "At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide." Proceedings of the National Academy of Sciences of the United States **90**: 8234-8238.

38. Callea, F., C. Baronchelli, A. Rodolfi, M. Bonetti and M. Brisigotti (1995). "Histopathology of chronic viral hepatitis: Guidelines for a revised classification." Italian Journal of Gastroenterology **27**: 137-140.

39. Calne, R. (1993). Liver and pancreas transplantation. Clinical Aspects of Immunology Cambridge, Mass., USA, Blackwell Scientific Publications. 5th ed. 1753-1760.

40. Camps, J., N. Garcia, J. RizuBoj, M. Civeira and J. Prieto (1993). "Ribavirin in the treatment of chronic hepatitis C unresponsive to alfa interferon." Journal of Hepatology **19**: 408-412.

41. Carman, W., J. Korula, L. Wallace, R. MacPhee, L. Mimms and R. Decker (1995). "Fulminant reactivation of hepatitis B due to envelope protein mutant that escaped detection by monoclonal HBsAg ELISA." Lancet **345**: 1406-1407.

42. Carman, W., H. Thomas and E. Domingo (1993a). "Viral genetic variation: Hepatitis B virus as a clinical example." Lancet **341**: 349-353.

43. Carman, W., H. Thomas, A. Zuckerman and T. Harrison (1993b). 7. Hepatitis B Virus: Molecular variants. Viral Hepatitis: Scientific Basis and Clinical Management. Edinburgh, Churchill Livingstone. 115-136.

44. Casey, J., T. Brown, E. Colan, F. Wignall and J. Gerin (1993). "A genotype of hepatitis D virus that occurs in northern South America." Proceedings of the National Academy of Sciences of the United States **90**: 9016-9020.

45. Cha, T., E. Beall, B. Irvine, J. Kolberg, D. Chien, G. Kuo and M. Urdea (1992). "Use of a signature nucleotide sequence of hepatitis C virus for the detection of viral RNA in human serum and plasma." Proceedings of the National Academy of Sciences of the United States **89:** 7144-7148.

46. Chan, S., F. McOmish, E. Holmes, B. Dow, J. Peutherer, E. Follett, P. Yap and P. Simmonds (1992). "Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants." Journal of General Virology **73** 1131-1141.

47. de Chastonay, J. and G. Siegl (1987). "Replicative events in hepatitis A virus infected MRC-5 cells. Virology **157:** 268-275.

48. Chemello, L., L. Cavalletto, F. Noventa, P. Bonetti, C. Casarin, E. Bernardinello, P. Pontisso, C. Donada, P. Casarin, F. Belussi, M. Frezza and A. Alberti (1995). "Predictors of sustained response, relapse and no response in patients with chronic hepatitis C treated with interferon-α." Journal of Viral Hepatitis **2**: 91-96.

49. Choo, Q., G. Kuo, R. Ralston, A. Weiner, D. Chien, G. Van Nest, J. Han, K. Berger, K. Thudium, C. Kuo, J. Kansopon, J. McFarland, A. Tabrizi, K. Ching, B. Moss, L. Cummins, M. Houghton and E. Muchmore (1994). "Vaccination of chimpanzees against infection by the hepatitis C virus." Proceedings of the National Academy of Sciences of the United States **91**: 1294-1298.

50. Choo, Q., G. Kuo, A. Weiner, L. Overby, D. Bradley and M. Houghton (1989). "Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome." Science **244**: 359-362.

51. Choo, Q., K. Richman, J. Han, K. Berger, C. Lee, C. Dong, C. Gallegos, D. Coit, A. MedinaSelby, P. Barr, A. Weiner, D. Bradley, G. Kuo and M. Houghton (1991). "Genetic organization and diversity of the hepatitis C virus." Proceedings of the National Academy of Sciences of the United States **88**: 2451-2455.

52. Choo, S., H. So, J. Cho and W. Ryu (1995). "Association of hepatitis C virus particles with immunoglobulin: A mechanism for persistent infection." Journal of General Virology **76**: 2337-2341.

53. Collett, M., D. Anderson and E. Retzel (1988). "Comparisons of the pestivirus bovine viral diarrhoea virus with members of the *Flaviviridae*." Journal of General Virology **69**: 2637-2643.

54. Collett, M., V. Moennig and M. Horzinek (1989). "Recent advances in pestivirus research." Journal of General Virology **70**: 253-266.

55. Collier, L. and J. Oxford (1993). Chapter 20.3. Human Virology. Oxford University Press. 254-265.

56. Coppola, R. and M. Rizzetto (1995). "Infectious hepatitis: An old protagonist revisited." FORUM - Trends in Experimental and Clinical Medicine **5**: 253-260.

57. Cornu, C., C. Brusselmans, A. Geubel, A. Ferrani and G. Cornu (1994). "Anti-HCV seroconversion in multitransfused and immunocompromised patients. A long-term longitudinal study." Vox Sanguinis **67**: 286-290.

58. Cronberg, S., B. Hansson, M. Thermos and a. et (1984). "Hepatitis D (delta agent) in primary hepatocellular carcinoma and liver disease in Senegal." Liver **4**: 275-279.

59. Cullen, J., C. David, J. Wang, P. Becherer and S. Lemon (1995). "Subcellular distribution of large and small hepatitis delta antigen in hepatocytes of hepatitis delta virus superinfected woodchucks." Hepatology **22**: 1090-1100.

60. Dane, D., C. Cameron and M. Briggs (1970). "Virus-like particles in serum of patients with Australia-antigen-associated hepatitis." Lancet 695-698.

61. De Wit, R. and R. Coutinho (1983). "A patient with hepatitis due to delta agent in the Netherlands." Nederlands Tijdschrift voor Geneeskunde **127**: 1678-1680.

62. Deka, N., M. Sharma and R. Mukerjee (1994). "Isolation of the novel agent from human stool samples that is associated with sporadic non-A, non-B hepatitis." Journal of Virology **68**: 7810-7815.

63. Deubel, V., R. Kinney and D. Trent (1988). "Nucleotide sequence and deduced amino acid sequence of the nonstructural proteins of Dengue type 2 virus, Jamaica genotype: Comparative analysis of the full-length genome." Virology **165:** 234-244.

64. Diamond, D., G. Davis, K. Qian and J. Lau (1994). "Detection of hepatitis C viral sequences in formalin-fixed, paraffin-embedded liver tissue: Effect of interferon alpha therapy." Journal of Medical Virology **42**: 294-298.

65. Driesel, G., D. Wirth, K. Stark, R. Baumgarten, U. Sucker and E. Schreier (1994). "Hepatitis C virus (HCV) genotype distribution in German isolates: studies on the sequence variability in the E2 and NS5 region." Archives of Virology **139**: 379-388.

66. Duffy, L., F. Daum, E. Kahn, S. Teichberg, R. Pahwa, J. Fagin, K. Kenigsberg, M. Kaplan, S. Fisher and S. Pahwa (1986). "Hepatitis in children with acquired immune deficiency syndrome: Histopathologic and immunocytologic features." Gastroenterology **90:** 173-181.

67. Duvoux, C., J. Pawlotsky, A. Bastie, D. Cherqui, J. Metreau, J. Duval and D. Dhumeaux (1995). "HCV replication levels are low in end-stage HCV-related cirrhosis." Journal of Hepatology **23**(Suppl 1): 110.

68. Duvoux, C., J. Pawlotsky, D. Cherqui, M. Julien, J. Duval and D. Dhumeaux (1994). "Diagnosis of HCV recurrence after liver transplantation using branched DNA assay for HCV RNA quantitation." Transplantation **58**: 953-954.

69. Eckart, M., M. Selby, F. Masiarz, C. Lee, K. Berger, K. Crawford, C. Kuo, G. Kuo, M. Houghton and Q. Choo (1993). "The hepatitis C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor." Biochemical and Biophysical Research Communications **192**: 399-406.

70. Ekland, E. and D. Bartel (1995). "The secondary structure and sequence optimization of an RNA ligase ribozyme." Nucleic Acids Research **23:** 3231-3238.

71. Enomoto, N., M. Kurosaki, Y. Tanaka, F. Marumo and C. Sato (1994). "Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis." Journal of General Virology **75**: 1361-1369.

72. Farci, P., H. Alter, S. Govindarajan, D. Wong, R. Engle, R. Lesniewski, I. Mushahwar, S. Desai, R. Miller, N. Ogata and R. Purcell (1992). "Lack of protective immunity against reinfection with hepatitis C virus." Science **258**: 135-140.

73. Farges, O., C. Feray, D. Samuel, A. Krygier, M. Sebagh and H. Bismuth (1995). "The level of immunosuppression does not appear to influence the natural history of recurrent disease in patients transplanted for HCV cirrhosis." Journal of Hepatology **23**(Suppl 1): 84.

74. Feray, C., D. Samuel, V. Thiers, M. Gigou, F. Pichon, A. Bismuth, M. Reynes, P. Maisonneuve, H. Bismuth and C. Brechot (1992). "Reinfection of liver graft by hepatitis C virus after liver transplantation." Journal of Clinical Investigation **89**: 1361-1365.

75. Ferrari, C., A. Valli, L. Galati, A. Penna, P. Scaccaglia, T. Giuberti, C. Schianchi, G. Missale, M. Marin and F. Fiaccadori (1994). "T-cell response to structural and nonstructural hepatitis C virus antigens in persistent and self-limited hepatitis C virus infections." Hepatology **19**: 286-295.

76. Fong, T., M. Shindo, S. Feinstone, J. Hoofnagle and A. Di Bisceglie (1991). "Detection of replicative intermediates of hepatitis C viral RNA in liver and serum of patients with chronic hepatitis C." Journal of Clinical Investigation **88**: 1058-1060.

77. Garson, J., S. Brillanti, C. Ring, P. Perini, M. Miglioli and L. Barbara (1992). "Hepatitis C viraemia rebound after 'successful' interferon therapy in patients with chronic non-A, non-B hepatitis." Journal of Medical Virology **37**: 210-214.

78. Garson, J., M. Lenzi, C. Ring, F. Cassani, G. Ballardini, M. Briggs, R. Tedder and F. Bianchi (1991). "Hepatitis C viraemia in adults with type 2 autoimmune hepatitis." Journal of Medical Virology **34**: 223-226.

79. Genesca, J; R. Jardi, M. Buti et al (1987). "Hepatitis B virus replication in acute hepatitis B, acute hepatitis B virus-hepatitis delta virus coinfection and acute hepatitis delta superinfection." Hepatology **7**: 569-572.

79. Gerlich, W. (1993). 6. Hepatitis B virus: Structure and molecular virology. Viral Hepatitis: Scientific Basis and Clinical Management. Edinburgh, Churchill Livingstone. 83-113.

80. Govindarajan, S., K. Chin, A. Redeker and R. Peters (1984). "Fulminant B viral hepatitis: role of delta agent." Gastroenterology **86**: 1417-1420.

81. Grakoui, A., D. McCourt, C. Wychowski, S. Feinstone and C. Rice (1993a). "Characterization of the hepatitis C virus-encoded serine proteinase: Determination of proteinase-dependent polyprotein cleavage sites." Journal of Virology **67**(5): 2832-2843.

82. Grakoui, A., D. McCourt, C. Wychowski, S. Feinstone and C. Rice (1993b). "A second hepatitis C virus-encoded proteinase." Proceedings of the National Academy of Sciences of the United States **90**(22): 10583-10587.

83. Grakoui, A., C. Wychowski, C. Lin, S. Feinstone and C. Rice (1993c). "Expression and identification of hepatitis C virus polyprotein cleavage products." Journal of Virology **67**(3): 1385-1395.

84. Gretch, D., C. Bacchi, L. Corey, C. Dela Rosa, R. Lesniewski, K. Kowdley, A. Gown, I. Frank, J. Perkins and J. Carithers RL (1995). "Persistent hepatitis C virus infection after liver transplantation: Clinical and virological features." Hepatology **22**: 1-9.

85. Hasan, F., L. Jeffers, M. De Medina, K. Reddy, T. Parker, E. Schiff, M. Houghton, Q. Choo and G. Kuo (1990). "Hepatitis C-associated hepatocellular carcinoma." Hepatology **12**(3 I): 589-591.

86. He, L., D. Alling, T. Popkin, M. Shapiro, H. Alter and R. Purcell (1987). "Determining the size of non-A, non-B hepatitis virus by filtration." Journal of Infectious Diseases **156:** 636-340.

87. Heinz, F. (1992). "Comparative molecular biology of flaviviruses and hepatitis C virus." Arch Virol Suppl **4**: 163-171.

88. Herrero, J., J. Lasarte, J. Quiroga, B. Sangro, M. Betes, N. Garcia, M. Civeira, J. Cienfuegos, J. Hernandez, F. Pardo and J. Prieto (1995). "Quantification of liver HCV RNA in liver transplant (LT) recipients." Journal of Hepatology **23**(Suppl 1): 153.

89. Higashi, Y., S. Kakumu, K. Yoshioka, T. Wakita, M. Mizokami, K. Ohba, Y. Ito, T. Ishikawa, M. Takayanagi and Y. Nagai (1993). "Dynamics of genome change in the E2/NS1 region of hepatitis C virus in vivo." Virology **197**: 659-668.

90. Hijikata, M; H. Mizushima, Y. Tanji, Y. Komoda, Y. Hirowatari, Y; T. Akagi, N. Kato, K. Kimura, K. Shimotohno (1993). "Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus." Proceedings of the National Academy of Sciences of the United States of America **90**: 10773-10777

91. Hijikata, M., N. Kato, Y. Ootsuyama, M. Nakagawa, S. Ohkoshi and K. Shimotohno (1991). "Hypervariable regions in the putative glycoprotein of hepatitis C virus." Biochemical and Biophysical Research Communications **175**: 220-228.

92. Hino, K., S. Sainokami, K. Shimoda, H. Niwa and S. Iino (1994). "Clinical course of acute hepatitis C and changes in HCV markers." Digestive Diseases and Sciences **39**: 19-27.

93. Hohne, M., E. Schreier and M. Roggendorf (1994). "Sequence variability in the env-coding region of hepatitis C virus isolated from patients infected during a single source outbreak." Archives of Virology **137**: 25-34.

94. Holland, J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol and S. VandePol (1982). "Rapid evolution of RNA genomes." Science **215**: 1577-1585

95. Holmes, E., L. Zhang, P. Simmonds, C. Ludlam and A. Brown (1992). "Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient." Proceedings of the National Academy of Sciences of the United States **89**: 4835-4839.

96. Honda, M., S. Kaneko, A. Sakai, M. Unoura, S. Murakami and K. Kobayashi (1994). "Degree of diversity of hepatitis C virus quasispecies and progression of liver disease." Hepatology **20**: 1144-1151.

97. Hoofnagle, J., K. Mullen, D. Jones, V. Rustgi, A. Di Bisceglie, M. Peters, J. Waggoner, Y. Park and A. Jones (1986). "Treatment of chronic non-A, non-B hepatitis with recombinant human alpha interferon: a preliminary report." New England Journal of Medicine **315**: 1575-1578.

98. Hsu, H., M. Chang, C. Lee, K. Hsieh, Y. Ni, P. Chen and S. Chen D (1995). "Precore mutant of hepatitis B virus in childhood fulminant hepatitis B: An infrequent association." Journal of Infectious Diseases **171**: 776-781.

99. Hsu, H., T. Wright, S. Tsao, C. Combs, M. Donets, S. Feinstone and H. Greenberg (1994). "Antibody response to hepatitis C virus infection after liver transplantation." American Journal of Gastroenterology **89**: 1169-1174.

100. Hu, K., C. Yu and J. Vierling (1993). "One-step RNA polymerase chain reaction for detection of hepatitis C virus RNA." Hepatology **18**: 270-274.

101. Hulst, M., D. Westra, G. Wensvoort and R. Moormann (1993). "Glycoprotein E1 of hog cholera virus expressed in insect cells protects swine from hog cholera." Journal of Virology **67**: 5435-5442.

102. Humphrey, C., A. Andjaparidze, E. Cook Jr and D. Bradley (1988). Identification of enterically transmitted non-A, non-B hepatitis virus particles by solid-phase immune electron microscopy. Viral Hepatitis and Liver Disease. USA, Alan R. Liss, Inc. 148-151.

103. Innis, M., D. Gelfand, J. Sninsky and T. White, Ed. (1990). PCR protocols: a guide to methods and applications. San Diego, Academic Press, Inc.

104. Inoue, Y., K. Takeuchi, W. Chou, T. Unayama, K. Takahashi, I. Saito and T. Miyamura (1992). "Silent mother-to-child transmission of hepatitis C virus through two generations determined by comparative nucleotide sequence analysis of the viral cDNA." Journal of Infectious Diseases **166**: 1425-1428.

105. Izraeli, S., C. Pfleiderer and T. Lion (1991). "Detection of gene expression by PCR amplification of RNA derived from frozen heparinized whole blood." Nucleic Acids Research **19**: 6051.

106. Jawetz, E., J. Melnick and E. Adelberg (1974). Chapter 27: General properties of viruses. Review of Medical Microbiology. Japan, Lange Medical Publications & Maruzen Company, Limited. 11th ed. 292-329.

107. JianRen, G., C. YuanChing, J. HuiQiu and a. et (1985). "State of hepatitis B virus DNA in leucocytes of hepatitis B patients." Journal of Medical Virology **17**: 73-81.

108. Jilg, W, M. Roggendorf and F. Deinhardt (1985). "What's new in hepatitis delta virus?" Pathology Research and Practice **180**: 431-436

109. Jilg, W., M. Roggendorf, G. Zoulek and a. et (1984). "HEPATITIS B UND DELTA-AGENS. Hepatitis B and delta agent." Fortschritte der Medizin **102**: 29-33.

110. Kahn, E., A. Greco, F. Daum, M. Magid, R. Morecki, V. Mahnovski and V. Anderson. "Hepatic pathology in paediatric acquired immunodeficiency syndrome." Human Pathology **22**: 1111-1119.

111. Kanazawa, Y., N. Hayashi, E. Mita, T. Li, H. Hagiwara, A. Kasahara, H. Fusamoto and T. Kamada (1994). "Influence of viral quasispecies on effectiveness of interferon therapy in chronic hepatitis C patients." Hepatology **20**: 1121-1130.

112. Kao, J., P. Chen, M. Lai and D. Chen (1993). "Superinfection of heterologous hepatitis C virus in a patient with chronic type C hepatitis." Gastroenterology **105**: 583-587.

113. Kao, J., P. Chen, M. Lai, T. Wang and D. Chen (1995). "Quasispecies of hepatitis C virus and genetic drift of the hypervariable region in chronic type C hepatitis." Journal of Infectious Diseases **172**: 261-264.

114. Kaplan, P., R. Greenman, J. Gerin, R. Purcell and W. Robinson (1973). "DNA polymerase associated with human hepatitis B antigen." Journal of Virology **12**: 995-1005.

115. Karthigesu, V., L. Allison, M. Fortuin, M. Mendy, H. Whittle and C. Howard (1994). "A novel hepatitis B virus variant in the sera of immunized children." Journal of General Virology **75**: 443-448.

116. Kato, N., Y. Ootsuyama, H. Sekiya, S. Ohkoshi, T. Nakazawa, M. Hijikata and K. Shimotohno (1994). "Genetic drift in hypervariable region 1 of the viral genome in persistent hepatitis C virus infection." Journal of Virology **68**: 4776-4784.

117. Keese, P. and R. Symons (1985). "Domains in viroids: Evidence of intermolecular RNA rearrangements and their contribution to viroid evolution." Proceedings of the National Academy of Sciences of the United States **82:** 4582-4586.

118. Keeton, W. and J. Gould (1986). Biological Science. New York, London, W. W. Norton & Company.

119. Kita, H., T. Moriyama, T. Kaneko, I. Harase, M. Nomura, H. Miura, I. Nakamura, Y. Yazaki and M. Imawari (1993). "HLA B44-restricted cytotoxic T lymphocytes recognizing an epitope on hepatitis C virus nucleocapsid protein." Hepatology **18**(5): 1039-1044.

120. Kiyosawa, K., T. Sodeyama, E. Tanaka, Y. Nakano, S. Furuta, K. Nishioka, R. Purcell and H. Alter (1991). "Hepatitis C in hospital employees with needlestick injuries." Annals of Internal Medicine **115**: 367-369.

121. Klenerman, P., U. Meier, R. Phillips and A. McMichael (1995). "The effects of natural altered peptide ligands on the whole blood cytotoxic T lymphocyte response to human immunodeficiency virus." European Journal of Immunology **25**: 1927-1931.

122. Ko, Y., M. Ho, T. Chiang, S. Chang and P. Chang (1992). "Tattooing as a risk of hepatitis c virus infection." Journal of Medical Virology **38**: 288-291.

123. Koizumi, K., N. Enomoto, M. Kurosaki, T. Murakami, N. Izumi, F. Marumo and C. Sato (1995). "Diversity of quasispecies in various disease stages of chronic hepatitis C virus infection and its significance in interferon treatment." Hepatology **22**: 30-35.

124. Konig, V., J. Bauditz, H. Lobeck, R. Lusebrink, P. Neuhaus, G. Blumhardt, W. Bechstein, R. Neuhaus, R. Steffen and U. Hopf (1992). "Hepatitis C virus reinfection in allografts after orthotopic liver transplantation." Hepatology **16**: 1137-1143.

125. Koonin, E. (1991). "The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses." Journal of General Virology **72**: 2197-2206.

126. Koonin, E. and V. Dolja (1993). "Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences." Critical Reviews in Biochemistry and Molecular Biology **28**: 375-430.

127. Kos, A., R. Dijkema, A. Amberg, P. van der Meide and H. Schellekens (1986). "The hepatitis delta virus possesses a circular RNA." Nature **323**: 558-560.

128. Kosai, K., M. Kage and M. Kojiro (1991). "Clinicopathological study of liver involvement in cytomegalovirus infection in infant autopsy case." Journal of Gastroenterology and Hepatology **6**: 603-608.

129. Koziel, M., D. Dudley, N. Afdhal, Q. Choo, M. Houghton, R. Ralston and B. Walker (1993). "Hepatitis C virus (HCV)-specific cytotoxic T lymphocytes recognize epitopes in the core and envelope proteins of HCV." Journal of Virology **67**: 7522-7532.

130. Kremsdorf, D., C. Porchon, J. Kim, G. Reyes and C. Brechot (1991). "Partial nucleotide sequence analysis of a French hepatitis C virus: Implications for HCV genetic variability in the E2/NS1 Protein." Journal of General Virology **72**: 2557-2561.

131. Kumar, U., J. Monjardino and H. Thomas (1994). "Hypervariable region of hepatitis C virus envelope glycoprotein (E2/NS1) in an agammaglobulinemic patient." Gastroenterology **106**: 1072-1075.

132. Kurosaki, M., N. Enomoto, F. Marumo and C. Sato (1993). "Rapid sequence variation of the hypervariable region of hepatitis C virus during the course of chronic infection." Hepatology **18**: 1293-1299.

133. Kurosaki, M., N. Enomoto, F. Marumo and C. Sato (1994). "Evolution and selection of hepatitis C virus variants in patients with chronic hepatitis C." Virology **205**: 161-169.

134. Kuo, D., Q. Choo, H. Alter, G. Gitnick, A. Redeker, R. Purcell, T. Miyamura, D. Dienstag, M. Alter, C. Stevens, G. Tegtmeier, F. Bonino, M. Colombo, W. Lee, C. Kuo, K. Berger, J. Shuster, L. Overby, D. Bradley and M. Houghton (1989). "An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis." Science **44:** 362-364.

135. Lai, M. (1995). "Molecular biologic and pathogenetic analysis of hepatitis delta virus." Journal of Hepatology, Supplement **22**: 127-131.

136. Lampertico, P., A. Aroldi, G. Lunghi, A. Archenti, M. Rumi, R. Soffredini, A. Tarantino, C. Ponticelli and M. Colombo (1995). "The natural history of hepatitis C in renal allograft recipients: a 5-year prospective study of 443 patients." Journal of Hepatology **23**(Suppl 1): 174.

137. Lau, J., G. Davis, M. Brunson, K. Qian, H. Lin, S. Quan, R. DiNello, A. Polito and J. Scornik (1993). "Hepatitis C virus infection in kidney transplant recipients." Hepatology **18**: 1027-1031.

138. Laugaa, P., A. Woisard, J. Fourrey and A. Favre (1995). "RIBOZYME EN TETE DE MARTEAU: UN MODELE TRIDIMENSIONNEL DEDUIT DE DONNEES DE PHOTOPONTAGE. Hammerhead ribozyme: A three dimentional model based on photocrosslinking data". Comptes Rendus de L'Academie des Sciences - Serie III **318:** 307-313.

139. Lebovics, E., B. Dworkin and S. Heier (1988). "The hepatobiliary manifestations of human immunodeficiency virus infections." American Journal of Gastroenterology **83:** 1-7.

140. Lemon, S., R. Jansen and E. Brown (1992). "Genetic, antigenic and biological differences between strains of hepatitis A virus." Vaccine **10**(S1): S40-S44.

141. Lerut, J., M. Donataccio, O. Cicarelli, J. Bail, C. Cornu, P. Laterre, F. Roggen, A. Geubel, J. Rahier and J. Otte (1995). "Liver transplantation and hepatitis B, C and delta viral diseases." Transplantationsmedizin: Organ der Deutschen 7: 41-46.

142. Linke, H., C. Amann and G. Schlauder (1988). Non-A, non-B hepatitis infection does not result in the production of abundant poly-A-containing messenger RNAs. Viral Hepatitis and Liver Disease. USA, Alan R. Liss, Inc. 564-567.

143. Linnen, J., J. Wages Jr, Z. ZhangKeck, K. Fry, K. Krawczynski, H. Alter, E. Koonin, M. Gallagher, M. Alter, S. Hadziyannis, P. Karayiannis, K. Fung, Y. Nakatsuji, J. Shih, L. Young, M. Piatak Jr, C. Hoover, J. Fernandez, J. Kim and a. et (1996). "Molecular cloning and disease association of hepatitis G virus: A transfusion-transmissible agent." Science **271**: 505-508.

144. Liou, T., T. Chang, K. Young, X. Lin, C. Lin and H. Wu (1992). "Detection of HCV RNA in saliva, urine, seminal fluid, and ascites." Journal of Medical Virology **37**: 197-202.

145. Lok, A., D. Chien, Q. Choo, T. Chan, E. Chiu, I. Cheng, M. Houghton and G. Kuo (1993). "Antibody response to core, envelope and nonstructural hepatitis C virus antigens: Comparison of immunocompetent and immunosuppressed patients." Hepatology **18**: 497-502.

146. Longo, M., M. Berninger and J. Hartley (1990). "Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions." Gene **93**: 125-128.

147. Lucey, M. (1995). "Interferons and liver disease." Saudi Medical Journal **16**: 83-86.

148. Lyngstadaas, S., S. Risnes, B. Sproat, P. Thrane and H. Prydz (1995). "A synthetic, chemically modified ribozyme eliminates amelogenenin, the major translation product in developing mouse enamel in vivo." EMBO Journal **14**: 5224-5229.

149. MacNaughton, T., Y. Wang and M. Lai (1993). "Replication of hepatitis delta virus RNA: Effect of mutations of the autocatalytic cleavage sites." Journal of Virology **67**: 2228-2234.

150. Main, J. (1995). Future studies of combination therapy for chronic hepatitis C: optimizing response rates for each of the hepatitis C populations. EASL 30th conference, Copenhagen 19-23 August.

151. Maple, P., T. McKee, U. Desselberger and T. Wreghitt (1994). "Hepatitis C virus infections in transplant patients: Serological and virological investigations." Journal of Medical Virology **44**(1): 43-48.

152. Marin, M., G. Ruggeri, N. Portolani, S. Bonardelli, F. Benzi, A. Rossini and E. Cariani (1995). "Hepatitis C virus (HCV) genotypes 1 and 2 in anti-HCV positive primary liver cancer (PLC)." Journal of Hepatology **23**(Suppl 1): 227.

153. Martell, M., J. Esteban, J. Quer, J. Genesca, A. Weiner, R. Esteban, J. Guardia and J. Gomez (1992). "Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: Quasispecies nature of HCV genome distribution." Journal of Virology **66**: 3225-3229.

154. Martell, M., J. Esteban, J. Quer, V. Vargas, R. Esteban, J. Guardia and J. Gomez (1994). "Dynamic behavior of hepatitis C virus quasispecies in patients undergoing orthotopic liver transplantation." Journal of Virology **68**: 3425-3436.

155. Martin, P., L. Friedman and J. Dienstag (1993). 26. Diagnosis of Viral Hepatitis: Diagnostic approach. Viral Hepatitis: Scientific Basis and Clinical Management. Edinburgh, Churchill Livingstone. 393-409.

156. Martin, P., S. Munoz, A. Di Bisceglie, R. Rubin, J. Waggoner, V. Armenti, M. Moritz, B. Jarrell and W. Maddrey (1991). "Reccurrence of hepatitis C virus infection after orthotopic liver transplantation." Hepatology **13**: 719-721.

157. Mason, W. and J. Taylor (1989). "Experimental system for the study of hepadnavirus and hepatitis delta virus infections." Hepatology **9**: 635-645.

158. Matsuura, Y., S. Harada, R. Suzuki, Y. Watanabe, Y. Inoue, I. Saito and T. Miyamura (1992). "Expression of processed envelope protein of hepatitis C virus in mammalian and insect cells." Journal of Virology **66**: 1425-1431.

159. McAdam, S., P. Klenerman, L. Tussey, S. RowlandJones, D. Lalloo, R. Phillips, A. Edwards, P. Giangrande, A. Brown, F. Gotch and A. McMichael (1995). "Immunogenic HIV variant peptides that bind to HLA-B8 can fail to stimulate cytotoxic T lymphocyte responses." Journal of Immunology **155**: 2729-2736.

160. McCuskey, R. (1994). 57: The hepatic microvascular system. The Liver: Biology and Pathobiology. New York, Raven Press Ltd. 2nd ed. 1089-1106.

161. McMichael, A., P. Klenerman, S. RowlandJones, F. Gotch and P. Moss (1994). "Recognition of viral antigens at the cell surface." Cancer Surveys **22**: 51-62.

162. Melbye, M., R. Biggar, P. Wantzin, K. Krogsgaard, P. Ebbesen and N. Becker (1990). "Sexual transmission of hepatitis C virus: Cohort study (1981-9) among European homosexual men." British Medical Journal **301**: 210-212.

163. Michel, G., A. Ritter, G. Gerken, K. Meyer zum Buschenfelde, R. Decker and M. Manns (1992). "Anti-GOR and hepatitis C virus in autoimmune liver diseases." Lancet **339**: 267-269.

164. Miller, R. and R. Purcell (1990). "Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups." Proceedings of the National Academy of Sciences of the United States **87**: 2057-2061.

165. Minor, P. (1991). Classification and nomenclature of viruses: Fifth Report on the International Committee on Taxonomy of Viruses. Archives of Virology (suppl 2). Springer Verlag, Vienna. 320-326.

166. Miyamoto, H., H. Okamoto, K. Sato, T. Tanaka and S. Mishiro (1992). "Extraordinarily low density of HCV estimated by surcrose density gradient centrifugation and the polymerase chain reaction." Journal of General Virology **73**: 715-718.

167. Muller, H., E. Pfaff, T. Goeser, B. Kallinowski, C. Solbach and L. Theilmann (1993). "Peripheral blood leukocytes serve as a possible extrahepatic site for hepatitis C virus replication." Journal of General Virology **74**: 669-676.

168. Murphy, F. and D. Kingsbury (1991). Chapter 2: Virus Taxonomy. Fundamental Virology. New York, Raven Press, Ltd. 9-35.

169. Navas, S., I. Castillo, J. Bartolome, E. Marriott, M. Herrero and V. Carreno (1994). "Positive and negative hepatitis C virus RNA strands in serum, liver and peripheral blood mononuclear cells in anti-HCV patients. Relation with the liver lesion." Journal of Hepatology **21**: 182-186.

170. Negro, F. and M. Rizzetto (1995). "Diagnosis of hepatitis delta virus infection." Journal of Hepatology, Supplement **22**: 136-139.

171. Netter, H., T. Wu, M. Bockol, A. Cywinski, W. Ryu, B. Tennant and J. Taylor (1995). "Nucleotide sequence stability of the genome of hepatitis delta virus." Journal of Virology **69**: 1687-1692.

172. Ng, Y., S. Lee, S. Wu, W. Yang, S. Chiang and T. Huang (1991). "Antibodies to hepatitis C virus in uremic patients on continuous ambulatory peritoneal dialysis." Journal of Medical Virology **35**: 263-266.

173. Nolte, F., C. Thurmond and P. Mitchell (1994). "Isolation of hepatitis C virus RNA from serum for reverse transcription- PCR." Journal of Clinical Microbiology **32**: 519-520.

174. Numata, N., H. Ohori, Y. Hayakawa, Y. Saitoh, A. Tsunoda and A. Kanno (1993). "Demonstration of hepatitis C virus genome in saliva and urine of patients with type C hepatitis: Usefulness of the single round polymerase chain reaction method for detection of the HCV genome." Journal of Medical Virology **41**: 120-128.

175. Ogata, N., H. Alter, R. Miller and R. Purcell (1991). "Nucleotide sequence and mutation rate of the H strain of hepatitis C virus." Proceedings of the National Academy of Sciences of the United States **88**: 3392-3396.

176. Ohkawa, J., T. Koguma, T. Kohda and K. Taira (1995). "Ribozymes: From mechanistic studies to applications in vivo." Journal of Biochemistry **118**: 251-258.

177. Okamoto, H., M. Kojima, M. Sakamoto, H. Iizuka, S. Hadiwandowo, S. Suwignyo, Y. Miyakawa and M. Mayumi (1994a). "The entire nucleotide sequence and classification of a hepatitis C virus isolate of a novel genotype from an Indonesian patient with chronic liver disease." Journal of General Virology **75**: 629-635.

178. Okamoto, H., S. Mishiro, H. Tokita, F. Tsuda, Y. Miyakawa and M. Mayumi (1994b). "Superinfection of chimpanzees carrying hepatitis C virus of genotype II/1b with that of genotype III/2a or I/1a." Hepatology **20**: 1131-1136.

179. Okamoto, H., H. Tokita, M. Sakamoto, M. Horikita, M. Kojima, H. Iizukas and S. Mishiro (1993). "Characterization of the genomic sequence of type V (or 3a) hepatitis C virus isolates and PCR primers for specific detection." Journal of General Virology **74**: 2385-2390.

180. Okamoto, H., K. Kurai, S. Okada, K. Yamamoto, H. Lizuka, T. Tanaka, S. Fukuda, F. Tsuda and S. Mishiro (1992). "Full-length sequence of a hepatitis C virus genome having poor

homology to reported isolates: Comparative study of four distinct genotypes." Virology **188**: 331-341.

181. Okano, M., S. Matsumoto, T. Osato, Y. Sakiyama, G. Thiele and D. Purtillo (1991). "Severe chronic active Epstein-Barr virus infection syndrome." Clinical Microbiology Review **4**:129-135.

182. Orito, E., M. Mizokami, N. Mizoguchi, K. Ohba, M. Tohnai, H. Yamanaka, T. Oguri, N. Hirashima, T. Koide, H. Kano, K. TsukiyamaKohara, M. Kohara and J. Lau (1994). "Hepatitis C virus serotype II responds more favorably to interferon-alpha therapy." Journal of Hepatology **21**: 130-132.

183. Ouanian, A., N. Gueddah, A. Rolachon, M. Thelu, J. Zarski and J. Seigneurin (1995). "Hepatitis C virus RNA in plasma and blood mononuclear cells in patients with chronic hepatitis C treated with alpha-interferon." Journal of Medical Virology **45**: 141-145.

184. Paton, A. (1969). Liver Disease. London, William Heinemann Medical Books Ltd.

185. Pereira, B., T. Wright, C. Schmid and A. Levey (1995). "A controlled study of hepatitis C transmission by organ transplantation." Lancet **345**: 469-470+484-487.

186. Petrik, J., G. Pearson and J-P Allain (1996). "High throughput PCR detection of HCV and HGV based on semiautomated multisample RNA capture" - submitted for publication

187. Perlman, D. and H. Halvorson (1983). "A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides." Journal of Molecular Biology **167**: 391-409.

188. Plagemann, P. (1991). "Hepatitis C virus: Brief review." Archives of Virology **120**: 165-180.

189. Poisson, F., P. Roingeard, A. Baillou, F. Dubois, F. Bonelli, R. Calogero and A. Goudeau (1993). "Characterization of RNA-binding domains of hepatitis delta antigen." Journal of General Virology **74**: 2473-2478.

190. Polo, J., K. Jeng, B. Lim, S. Govindarajan, F. Hofman, F. Sangiorgi and M. Lai (1995). "Transgenic mice support replication of hepatitis delta virus RNA in multiple tissues, particularly in skeletal muscle." Journal of Virology **69**: 4880-4887.

191. Pons, J. (1995). "Role of liver transplantation in viral hepatitis." Journal of Hepatology, Supplement **22**(1): 146-153
Conference.

192. Poterucha, J., J. Rakela, L. Lumeng, C. Lee, H. Taswell and R. Wiesner (1992). "Diagnosis of chronic hepatitis C after liver transplantation by the detection of viral sequences with polymerase chain reaction." Hepatology **15**: 42-45.

193. Pozzato, G., S. Kaneko, M. Moretti, L. Croce, F. Franzin, M. Unoura, L. Bercich, C. Tiribelli, M. Crovatto, G. Santini and K. Kobayashi (1994). "Different genotypes of hepatitis C virus are associated with different severity of chronic liver disease." Journal of Medical Virology **43**: 291-296.

194. Prince, A., B. Brotman, T. Huima, D. Pascual, M. Jaffery and G. Inchauspe (1992). "Immunity in hepatitis C infection." Journal of Infectious Diseases **165**: 438-443.

195. Provost, P. and M. Hilleman (1975). "An inactivated hepatitis A virus vaccine prepared from infected marmoset liver." Proceedings of the Society for Experimental Biology and Medicine **159**: 201-213.

196. Purcell, R. and J. Ticehurst (1988). Enterically transmitted non-A, non-B hepatitis: Epidemiology and clinical characteristics. Viral Hepatitis and Liver Disease. USA, Alan R. Liss, Inc. 131-137.

197. Purdy, M. and K. Krawczynski (1994). "Hepatitis E." Gastroenterology Clinics of North America **23**: 537-546.

198. Ralston, R., K. Thudium, K. Berger, C. Kuo, B. Gervase, J. Hall, M. Selby, G. Kuo, M. Houghton and Q. Choo (1993). "Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses." Journal of Virology **67**: 6753-6761.

199. Read, R., R. Ede, P. Morgan-Capner, G. Moscoso, B. Portmann and R. Williams (1985). "Myocarditis and fulminant hepatic failure from coxsackie virus B infection." Postgraduate Medical Journal **61**: 749-752.

200. Rees, A. and C. Lockwood (1993). Immunosuppressive drugs in clinical practice. Clinical Aspects of Immunology. Cambridge, Mass., USA, Blackwell Scientific Publications. 5th ed. 929-972.

201. Reichard, O., J. Andersson, R. Schvarcz and O. Weiland (1991). "Ribavirin treatment for chronic hepatitis C." Lancet **337**: 1058-1061.

202. Reyes, G., M. Purdy, J. Kim, K. Luk, L. Young, K. Fry and D. Bradley (1990). "Isolation of a cDNA from the virus responsible for enterically transmitted non-A, non-B hepatitis." Science **247**: 1335-1339.

203. Rice, C., E. Lenches, S. Eddy, S. Shin, R. Sheets and J. Strauss (1985). "Nucleotide sequence of yellow fever virus: Implications for flavivirus gene expression and evolution." Science **229**: 726-733.

204. Riesner, D., K. Henco, U. Rokohl, G. Klotz, A. Kleinschmidt, H. Domdey, P. Jank, H. Gross and H. Sanger (1979). "Structure and structure formation of viroids." Journal of Molecular Biology **133**: 85-115.

205. Rizzetto, M. (1983). "The delta agent." Hepatology **3**: 729-737.

206. Robertson, B., B. Khanna, O. Nainan and H. Margolis (1991). "Epidemiologic patterns of wild-type hepatitis A virus determined by genetic variation." Journal of Infectious Diseases **163**: 286-292.

207. Robinson, W., D. Clayton and R. Greenman (1974b). "DNA of a human hepatitis B virus candidate." Journal of Virology **14**: 384-391.

208. Robinson, W. and R. Greenman (1974a). "DNA polymerase in the core of the human hepatitis B virus candidate." Journal of Virology **13**: 1231-1236.

209. Rojkind, M. and P. Greenwel (1994). 43: The extracellular matrix of the liver. The Liver: Biology and Pathobiology. New York, Raven Press Ltd: 2nd ed. 843-868.

210. Rosa, D., S. Campagnoli, C. Moretto, E. Guenzi, L. Cousens, M. Chin, C. Dong, A. Weiner, J. Lau, Q.-L. Choo, D. Chien, P. Pileri, M. Houghton and S. Abrignani (1996). "A quantitative test to estimate neutralizing antibodies to the hepatitis C virus: Cytofluorimetric assessment of envelope glycoprotein 2 binding to target cells." Proceedings of the National Academy of Science **93**: 1759-1763.

211. RowlandJones, S., J. Sutton, K. Ariyoshi, T. Dong, F. Gotch, S. McAdam, D. Whitby, S. Sabally, A. Gallimore, T. Corrah, M. Takiguchi, T. Schultz, A. McMichael and H. Whittle

(1995). "HIV-specific cytotoxic T-cells in HIV-exposed but uninfected Gambian women." Nature Medicine **1**(1): 59-64.

212. Saldanha, J. and P. Minor (1994). "A sensitive PCR method for detecting HCV RNA in plasma pools, blood products, and single donations." Journal of Medical Virology **43**: 72-76.

213. Saldanha, J., H. Thomas and J. Monjardino (1990). "Cloning and sequencing of RNA of hepatitis delta virus isolated from human serum." Journal of General Virology **71:** 1603-1606.

214. Sallie, R., A. Cohen, C. Tibbs, B. Portmann, A. Rayner, J. OGrady, K. Tan and R. Williams (1994). "Recurrence of hepatitis C following orthotopic liver transplantation: A polymerase chain reaction and histological study." Journal of Hepatology **21**: 536-542.

215. Samuel, D., A. Zignego, M. Reynes, C. Feray, J. Arulnaden, F. David M, M. Gigou, A. Bismuth, D. Mathieu, P. Gentilini, J. Benhamou, C. Brechot and H. Bismuth (1995). "Long-term clinical and virological outcome after liver transplantation for cirrhosis caused by chronic delta hepatitis." Hepatology **21**: 333-339.

216. Satsangi, J., D. Jewell, K. Welsh, M. Bunce and J. Bell (1994). "Effect of heparin on polymerase chain reaction (19)." Lancet **343**: 1509-1510.

217. Schlauder, G., G. Dawson, J. Simons, T. PilotMatias, G. Gutierrez, C. Heynen, M. Knigge, G. Kurpiewski, S. Buijk, T. Leary, A. Muerhoff, S. Desai and I. Mushahwar (1995). "Molecular and serologic analysis in the transmission of the GB hepatitis agents." Journal of Medical Virology **46**: 81-90.

218. Schneiderman, D., D. Arenson, J. Cello, W. Margaretten and T. Weber (1987). "Hepatic disease in patients with the acquired immune deficiency syndrome (AIDS)." Hepatology **7:** 925-930.

219. Seeger, C., D. Ganem and H. Varmus (1986). "Biochemical and genetic evidence for the hepatitis B virus replication strategy." Science **232:** 477-483.

220. Selby, M., Q. Choo, K. Berger, G. Kuo, E. Glazer, M. Eckart, C. Lee, D. Chien, C. Kuo and M. Houghton (1993). "Expression, identification and subcellular localization of the proteins encoded by the hepatitis C viral genome." Journal of General Virology **74**(6): 1103-1113.

221. Sheiner, P., M. Schwartz, E. Mor, L. Schluger, N. Theise, K. Kishikawa, V. Kolesnikov, H. Bodenheimer, S. Emre and C. Miller (1995). "Severe or multiple rejection episodes are

associated with early recurrence of hepatitis C after orthotopic liver transplantation." Hepatology **21**: 30-34.

222. Sherlock, S. (1995). Combination antiviral therapy: a new paradigm for hepatitis C management. EASL 30th conference, Copenhagen 19-23 August.

223. Shiffman, M., M. Contos, V. Luketic, A. Sakval, I. Purdum PP, A. Mills, R. Fisher and M. Posner (1994). "Biochemical and histologic evaluation of recurrent hepatitis C following orthotopic liver transplantation." Transplantation **57**: 526-532.

224. Shimizu, Y., M. Hijikata, A. Iwamoto, H. Alter, R. Purcell and H. Yoshikura (1994). "Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses." Journal of Virology **68**: 1494-1500.

225. Shindo, M., A. Di Bisceglie, L. Cheung, J. Shih, K. Cristiano, S. Feinstone and J. Hoofnagle (1991). "Decrease in serum hepatitis C viral RNA during alpha-interferon therapy for chronic hepatitis C." Annals of Internal Medicine **115**: 700-704.

226. Shirai, M., T. Arichi, M. Nishioka, T. Nomura, K. Ikeda, K. Kawanishi, V. Engelhard, S. Feinstone and J. Berzofsky (1995). "CTL responses of HLA-A2.1-transgenic mice specific for hepatitis C viral peptides predict epitopes for CTL of humans carrying HLA-A2.1." Journal of Immunology **154**: 2733-2742

227. Shirai, M., H. Okada, M. Nishioka, T. Akatsuka, C. Wychowski, R. Houghten, C. Pendleton, S. Feinstone and J. Berzofsky (1994). "An epitope in hepatitis C virus core region recognized by cytotoxic T cells in mice and humans." Journal of Virology **68**: 3334-3342.

228. Siegl, G., J. de Chastonay and G. Kronauer (1984). "Propagation and assay of hepatitis A virus in vitro. Journal of Virological Methods **9**: 53-67.

229. Simmonds, P. (1995). "HCV genotypes - nomenclature." INTER-action **3**: 4-9.

230. Simmonds, P., E. Holmes, T. Cha, S. Chan, F. McOmish, B. Irvine, E. Beall, P. Yap, J Kolberg and M. Urdea (1993a). "Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region." Journal of General Virology **74**:2391-2399.

231. Simmonds, P., K. Rose, S. Graham, S.-W. Chan, F. McOmish, B. Dow, E. Follett, P. Yap and H. Marsden (1993b). "Mapping of serotype-specific, immunodominant epitopes in the NS-4

region of hepatitis C virus (HCV): Use of type-specific peptides to serologically differentiate infections with HCV types 1, 2 and 3." Journal of Clinical Microbiology **31**: 1493-1503.

232. Simons, J., T. Leary, G. Dawson, T. PilotMatias, A. Muerhoff, G. Schlauder, S. Desai and I. Mushahwar (1995). "Isolation of novel virus-like sequences associated with human hepatitis." Nature Medicine **1**: 564-569.

233. Simons, J., T. PilotMatias, T. Leary, G. Dawson, S. Desai, G. Schlauder, A. Muerhoff, J. Erker, S. Buijk, M. Chalmers, C. Van Sant and I. Mushahwar (1995). "Identification of two flavivirus-like genomes in the GB hepatitis agent." Proceedings of the National Academy of Sciences of the United States **92**: 3401-3405.

234. Smedile, A., G. Verme, A. Cargnel, P. Dentico, P. Opolon, D. Vergani, P. Farci, F. Caredda, N. Caporase, C. Trepo, A. Gimson, R. Williams and M. Rizzetto (1982). "Influence of delta infection on severity of hepatitis B." Lancet **2**: 945-947.

235. Sonigo, P. and P. Tiollais (1985). "LE VIRUS DE L'HEPATITE B. The hepatitis B virus." Bulletin de l'Institut Pasteur **83**: 181-206

236. Spaete, R., D. Alexander, M. Rugroden, Q. Choo, K. Berger, K. Crawford, C. Kuo, S. Leng, C. Lee, R. Ralston, K. Thudium, J. Tung, G. Kuo and M. Houghton (1992). "Characterization of the hepatitis C virus E2/NS1 gene product expressed in mammalian cells." Virology **188**: 819-830.

237. Stark, R., T. Rumenapf, G. Meyers and H. Thiel (1990). "Genomic localization of hog cholera virus glycoproteins." Virology **174**: 286-289.

238. Starzl, T., A. Demetris and D. Van Thiel (1989a). "Liver transplantation (first of two parts)." New England Journal of Medicine **321**: 1014-1022.

239. Starzl, T., A. Demetris and D. Van Thiel (1989b). "Liver transplantation (Second of two parts)." New England Journal of Medicine **321**: 1092-1099.

240. Strauss, E., J. Strauss and A. Levine (1991). Chapter 9: Virus Evolution. Fundamental Virology. New York, Raven Press, Ltd. 167-190.

241. Stroffolini, T., L. Ferrigno, L. Cialdea, R. Catapano, F. Palumbo, F. Novaco, A. Moiraghi, C. Galanti, R. Bernacchia and A. Mele (1994). "Incidence and risk factors of acute

Delta hepatitis in Italy: Results from a national surveillance system." Journal of Hepatology **21**: 1123-1126.

242. Sugaya, N., M. Nirasawa, K. Mitamura and A. Murata (1988). "Hepatitis in acquired rubella infection in children." American Journal of Diseases in Children **142:** 817-818.

243. Tabor, E. (1985). "The three viruses of non-A, non-B hepatitis." Lancet 1: 743-745.

244. Takamizawa, A., C. Mori, I. Fuke, S. Manabe, S. Murakami, J. Fujita, E. Onishi, T. Andoh, I. Yoshida and H. Okayama (1991). "Structure and organization of the hepatitis C virus genome isolated from human carriers." Journal of Virology **65**: 1105-1113.

245. Takehara, T., N. Hayashi, E. Mita, H. Hagiwara, K. Ueda, K. Katayama, A. Kasahara, H. Fusamoto and T. Kamada (1992). "Detection of the minus strand of hepatitis C virus RNA by reverse transcription and polymerase chain reaction: Implications for hepatitis C virus replication in infected tissue." Hepatology **15**: 387-390.

246. Taniguchi, S., H. Okamoto, M. Sakamoto, M. Kojima, F. Tsuda, T. Tanaka, E. Munekata, E. Muchmore, D. Peterson and S. Mishiro (1993). "A structurally flexible and antigenically variable N-terminal domain of the hepatitis C virus E2/NS1 protein: Implication for an escape from antibody." Virology **195**: 297-301.

247. Tedder, R., R. Gilson, M. Briggs, C. Loveday, C. Cameron, J. Garson, G. Kelly and I. Weller (1991). "Hepatitis C virus: Evidence for sexual transmission." British Medical Journal **302**: 1299-1302.

248. Tibbs, C. and R. Williams (1995). "Liver transplantation for acute and chronic viral hepatitis." Journal of Viral Hepatitis **2**: 65-72.

249. Tiollais, P., C. Pourcel and A. Dejean (1985). "The hepatitis B virus." Nature **317**: 489-495

250. TsukiyamaKohara, K., N. Iizuka, M. Kohara and A. Nomoto (1992). "Internal ribosome entry site within hepatitis C virus RNA." Journal of Virology **66**: 1476-1483.

251. Urdang, L. and H. Swallow, Ed. (1983). Mosby's Medical & Nursing Dictionary. St. Louis, Toronto, London, The C.V. Mosby Company.

252. Van der Poel, C., H. Cuypers, H. Reesink, Q. Choo, G. Kuo, J. Han, S. Quan, A. Polito, J. Verstraten, J. Van de Wouw, W. Schaasberg, M. Houghton and P. Lelie (1991). "Risk factors in hepatitis C virus-infected blood donors." Transfusion **31**: 777-779.

253. Van Doorn, L., I. Capriles, G. Maertens, R. DeLeys, K. Murray, T. Kos, H. Schellekens and W. Quint (1995). "Sequence evolution of the hypervariable region in the putative envelope region E2/NS1 of hepatitis C virus is correlated with specific humoral immune responses." Journal of Virology **69**: 773-778.

254. Van Doorn, L., W. Quint, K. Tsiquaye, J. Voermans, D. Paelinck, T. Kos, G. Maertens, H. Schellekens and K. Murray (1994). "Longitudinal analysis of hepatitis C virus infection and genetic drift of the hypervariable region." Journal of Infectious Diseases **169**: 1226-1235.

255. Vigevani, G., M. Almaviva, M. Ragni and a. et (1985). "CONSIDERAZIONI SU UNA CASISTICA DI EPATITE VIRALE FULMINANTE (EVF) IN RAPPORTO ALL'AGENTE DELTA. Does delta agent interfere with HBV antigens during fulminant viral hepatitis (FVH)?" Giornale di Malattie Infettive e Parassitarie **37**: 466-468.

256. von Heijne, G. (1985). "Signal sequences: the limits of variation." Journal of Molecular Biology **184**: 99-105.

257. Wallace, L., J. Echevarria, J. Echevarria and W. Carman (1994). "Molecular characterization of envelope antigenic variants of hepatitis B virus from Spain." Journal of Infectious Diseases **170**: 1300-1303.

258. Wang, C., P. Sarnow and A. Siddiqui (1993). "Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism." Journal of Virology **67**: 3338-3344.

259. Waters, J., M. Kennedy, P. Voet, P. Hauser, J. Petre, W. Carman and H. Thomas (1992). "Loss of the common 'A' determinant of hepatitis B surface antigen by a vaccine-induced escape mutant." Journal of Clinical Investigation **90**: 2543-2547.

260. Wattre, P. (1994). "LE VIRUS DE L'HEPATITE E. Hepatitis E virus: A review." Annales de Biologie Clinique **52**: 507-513.

261. Weiland, O., R. Schvarcz, Z. Yun and A. Sonnerborg (1995). IFN relapsers and nonresponders: Scandinavian experience. EASL 30th conference, Copenhagen 19-23 August.

262. Weiner, A., M. Thaler, K. Crawford, K. Ching, J. Kansopon, D. Chien, J. Hall, F. Hu and M. Houghton (1993). "A unique, predominant hepatitis C virus variant found in an infant born to a mother with multiple variants." Journal of Virology **67**: 4365-4368.

263. Weiner, A., H. Geysen, C. Christopherson, J. Hall, T. Mason, G. Saracco, F. Bonino, K. Crawford, C. Marion, K. Crawford, M. Brunetto, P. Barr, T. Miyamura, J. McHutchinson and M. Houghton (1992). "Evidence for immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: Potential role in chronic HCV infections." Proceedings of the National Academy of Sciences of the United States **89**: 3468-3472.

264. Weiner, A., M. Brauer, J. Rosenblatt, K. Richman, J. Tung, K. Crawford, F. Bonino, G. Saracco, Q. Choo, M. Houghton and J. Han (1991). "Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins." Virology **180**: 842-848.

265. Weitz, M. and G. Siegl (1993). 2. Hepatitis A virus: Structure and molecular virology. Viral Hepatitis: Scientific Basis and Clinical Management. Edinburgh, Churchill Livingstone. 21-34.

266. White, D. and F. Fenner (1986). Chapter 13. Hepadnaviruses. Medical Virology 3rd edition. Academic Press, Inc. 365-380. Orlando

267. White, D. and F. Fenner (1986). Medical Virology. Academic Press, Inc. 3rd ed. 607-632.

268. Willems, M., K. Peerlinck, H. Moshage, I. Deleu, C. Van den Eynde, J. Vermylen and S. Yap (1994). "Hepatitis C virus-RNAs in plasma and in peripheral blood mononuclear cells of hemophiliacs with chronic hepatitis C: Evidence for viral replication in peripheral blood mononuclear cells." Journal of Medical Virology **42**: 272-278.

269. Williams, K., S. Ciafre and G. TocchiniValentini (1995). "Selection of novel Mg^{2+}-dependent self-cleaving ribozymes." EMBO Journal **14**: 4551-4557.

270. Wong, V., D. Wight, C. Palmer and G. Alexander (1996). "Fibrosis and other histological features in chronic hepatitis C virus infection: a statistical model." Journal of Clinical Pathology **49**: 1-6.

271. Wreghitt, T., J. Gray, J. Allain, J. Poulain, J. Garson, R. Deaville, C. Maple, J. Parameshwar, R. Calne, J. Wallwork and G. Alexander (1994). "Transmission of hepatitis C virus by organ transplantation in the United Kingdom." Journal of Hepatology **20**: 768-772.

272. Wright, T., E. Donegan, H. Hsu, L. Ferrell, J. Lake, M. Kim, C. Combs, S. Fennessy, J. Roberts, N. Ascher and H. Greenberg (1992). "Recurrent and acquired hepatitis C viral infection in liver transplant recipients." Gastroenterology **103**: 317-322.

273. Wu, J., K. Choo, C. Chen, T. Chen, T. Huo and S. Lee (1995). "Genotyping of hepatitis D virus by restriction-fragment length polymorphism and relation to outcome of hepatitis D." Lancet **346**: 939-941.

274. Yamamoto, K., M. Horikita, F. Tsuda, K. Itoh, Y. Akahane, S. Yotsumoto, H. Okamoto, Y. Miyakawa and M. Mayumi (1994). "Naturally occurring escape mutants of hepatitis B virus with various mutations in the S gene in carriers seropositive for antibody to hepatitis B surface antigen." Journal of Virology **68**: 2671-2676.

275. Yoo, B., R. Spaete, A. Geballe, M. Selby, M. Houghton and J. Han (1992). "5' End-dependent translation initiation of hepatitis C viral RNA and the presence of putative positive and negative translationals control elements within the 5' untranslated region." Virology **191**: 889-899.

276. Yoshioka, K., S. Kakumu, T. Wakita, T. Ishikawa, Y. Itoh, M. Takayanagi, Y. Higashi, M. Shibata and T. Morishima (1992). "Detection of hepatitis C virus by polymerase chain reaction and response to interferon-alpha therapy: Relationship to genotypes of hepatitis C virus." Hepatology **16**: 293-299.

277. Zachoval, R. and F. Deinhardt (1993). 3. Hepatitis A virus: Natural history and experimental models. Viral Hepatitis: Scientific Basis and Clinical Management. Edinburgh, Churchill Livingstone. 35-53.

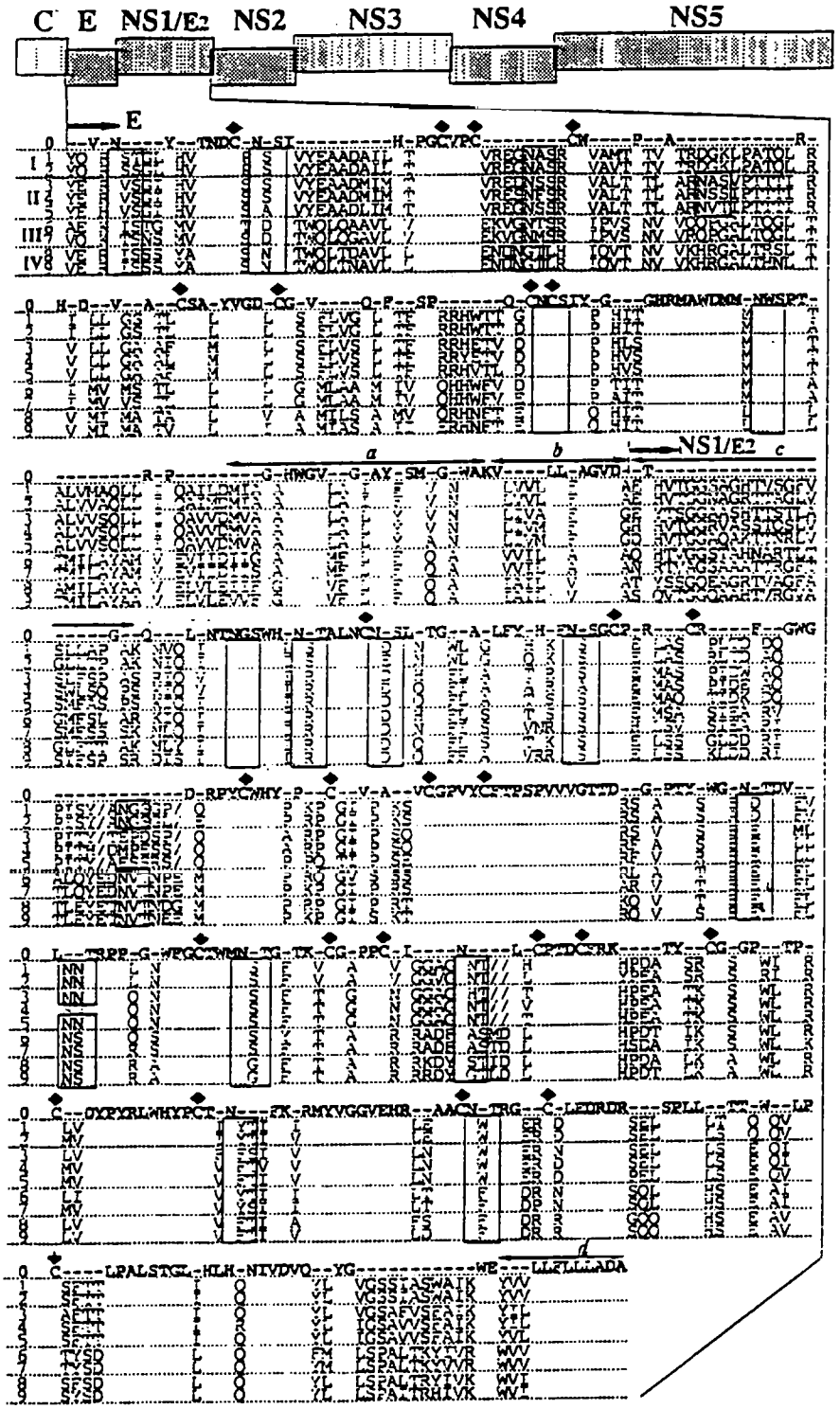
278. Zein, N., J. Gross, A. Gossard, R. Wiesner, J. Poterucha and D. Persing (1995). "Significance of hepatitis C (HCV) RNA levels in orthotopic liver transplant recipients (OLT)." Journal of Hepatology **23**(Suppl 1): 157.

279. Zhang, Z., A. Sonnerborg and M. Sallberg (1994). "Antigenic structure of the hepatitis C virus envelope 2 protein." Clinical and Experimental Immunology **98**: 382-387.

280. Zibert, A., E. Schreier and M. Roggendorf (1995). "Antibodies in human sera specific to hypervariable region 1 of hepatitis C virus can block viral attachment." Virology **208**: 653-661.

APPENDIX A

Region of HCV genome showing deduced amino acid sequences of putative E1 and E2 proteins from 9 HCV isolates representing 4 genotypes.



The relative degree of amino acid sequence variations was calculated for the four major isolates of HCV (HCV-1, HCV-J, HC-J6 and HC-J8. Sites having variation values of 3 or higher are shaded in the diagram shown at the top. Sequences 1-9 were from HCV-1, HCV-H, HC-J4, HCV-J, HCV-BK, HC-J6, HC-J5, HC-J8 and HC-J7, respectively. Conserved amino acids are shown in the sequence o. Putative N-glycosulation sites are boxed; conserved Cys residues are indicated by . A transmembtrane-like sequence is indicated by *a*, possible recognition sites for signal peptidase by *b* and *d*, and HVR1 by *c*. Down arrow indicates the beginning of the sequence used for type designation (amino acids 341 - 351).

Ref: Okamoto *et al.*, Virology **188**:338 (1992)

Appendix B:

Nucleotide and amino acid substitution tables for HCV-infected individuals in study.

Table .B.1 *Nucleotide (a) and amino acid (b)substitution tables for patient 1(Wad)*

*Boxed regions represent intraisolate sequence diversity. * T = no. of days since first sample; R = ratio of quasispecies; Q = designation of representative quasispecies.*

(a)

T*:		0			267
R*:	4:	4:	1	7:	4
Q*:	101	001	114	203	202
101	0	2	1	0	14
001	2	0	1	2	16
114	1	1	0	1	15
203	0	2	1	0	14
202	14	16	15	14	0

(b)

T*:		0			267
R*:	4:	4:	1	7:	4
Q*:	101	001	114	203	202
101	0	1	1	0	9
001	1	0	0	1	9
114	1	0	0	0	9
203	0	1	1	0	9
202	9	9	9	9	0

Table B.2 *Nucleotide (a) and amino acid (b)substitution tables for patient 2 (D1)*

*Boxed regions represent intraisolate sequence diversity. * T = no. of days since first sample; R = ratio of quasispecies; Q = designation of representative quasispecies.*

(a)

T*:	0		457				719									
R*:	3:	2	11:	1:	1:	1	2:	1:	1:	1:	1:	1:	1:	1:	1:	1
Q*:	111	113	268	265	266	267	301	302	303	304	305	307	316	318	319	320
111	0	11	14	4	13	14	12	14	13	14	3	5	11	3	16	10
113	11	0	13	16	10	13	11	13	12	12	14	10	10	10	13	12
268	14	13	0	18	11	1	4	13	13	10	17	15	11	15	15	15
265	4	15	18	0	15	18	16	17	16	16	5	8	15	7	19	10
266	13	10	11	15	0	11	11	15	15	12	14	12	12	12	17	10
267	14	13	1	18	11	0	4	13	13	12	17	15	11	15	15	15
301	12	11	4	16	11	4	0	11	11	10	15	13	9	13	13	13
302	14	13	13	17	15	13	11	0	1	6	17	14	5	14	4	13
303	13	12	13	16	15	13	11	1	0	6	16	13	5	13	4	14
304	14	12	10	16	12	12	10	6	6	0	17	13	5	13	8	15
305	3	14	17	5	14	17	15	17	16	17	0	8	14	6	19	9
307	5	10	15	8	12	15	13	14	13	13	8	0	11	2	16	7
316	11	10	11	15	12	11	9	5	5	5	14	11	0	11	5	11
318	3	10	15	7	12	15	13	14	13	13	6	2	11	0	16	8
319	16	13	15	19	17	15	13	4	4	8	19	16	5	16	0	16
320	10	12	15	10	10	15	13	13	14	15	9	7	11	8	16	0

(b)

T*:	0		457				719									
R*:	3:	2	11:	1:	1:	1	2:	1:	1:	1:	1:	1:	1:	1:	1:	1
Q*:	111	113	268	265	266	267	301	302	303	304	305	307	316	318	319	320
111	0	7	13	3	11	13	12	10	10	12	3	4	9	3	11	6
113	7	0	10	10	7	10	9	9	9	9	9	7	7	8	9	7
268	13	10	3	5	7	0	3	10	10	9	15	13	9	14	10	13
265	3	10	15	0	12	15	15	12	12	12	5	6	12	6	13	7
266	11	7	7	12	0	7	9	11	11	9	12	9	9	11	11	9
267	13	10	0	15	7	0	3	10	10	9	15	13	9	14	10	13
301	12	9	3	15	9	3	0	9	9	9	14	12	8	13	9	12
302	10	9	10	12	11	10	9	0	0	5	11	10	4	11	2	9
303	10	9	10	12	11	10	9	0	0	5	11	10	4	11	2	9
304	12	9	9	12	9	9	9	5	5	0	12	11	5	13	5	10
305	3	9	15	5	12	15	14	11	11	12	0	5	10	6	12	6
307	4	7	13	6	9	13	12	10	10	11	5	0	9	3	11	4
316	9	7	9	12	9	9	8	4	4	5	10	9	0	10	2	6
318	3	8	14	6	11	14	13	11	11	13	6	3	10	0	12	10
319	11	9	10	13	11	10	9	2	2	5	12	11	2	12	0	10
320	6	7	13	7	9	13	12	9	9	10	6	4	8	6	10	0

Table B.3 *Nucleotide (a) and amino acid (b)substitution tables for patient 3 (Fra)*

*Boxed regions represent intraisolate sequence diversity. * T = no. of days since first sample; R = ratio of quasispecies; Q = designation of representative quasispecies.*

(a)

T*:		0						24					87	
R*:	9:	1:	1	3:	2	2	1:	1:	1:	1:	1:	7:	5
Q*:	104	101	110	202	201	206	211	203	210	212	205	301	302
104	0	1	18	0	23	17	24	1	22	1	22	22	1
101	1	0	19	1	24	18	25	2	23	2	23	23	2
110	18	19	0	18	16	2	16	18	15	17	15	15	19
202	0	1	18	0	23	17	24	1	22	1	22	22	1
201	23	24	16	23	0	16	3	23	1	24	1	1	24
206	17	18	2	17	16	0	16	17	15	18	15	16	18
211	24	25	16	24	3	16	0	24	2	25	4	2	25
203	1	2	18	1	23	17	24	0	22	2	22	22	2
210	22	23	15	22	1	15	2	22	0	23	2	0	23
212	1	2	17	1	24	18	25	2	23	0	23	23	2
205	22	23	15	22	1	15	4	22	2	23	0	2	23
301	22	23	15	22	1	15	2	22	0	23	2	0	23
302	1	2	19	1	24	18	25	2	23	2	23	23	0

(b)

T*:		0						24					87	
R*:	9:	1:	1	3:	2	2	1:	1:	1:	1:	1:	7:	5
Q*:	104	101	110	202	201	206	211	203	210	212	205	301	302
104	0	1	13	0	16	13	18	0	16	0	15	16	0
101	1	0	14	1	17	14	19	1	17	1	16	17	1
110	13	14	0	13	12	0	12	13	11	13	11	11	13
202	0	1	13	0	16	13	18	0	16	0	15	16	0
201	16	17	12	16	0	12	3	16	1	16	1	1	16
206	13	14	0	13	12	0	12	13	11	13	11	11	13
211	18	19	12	18	3	12	0	18	2	18	4	2	18
203	0	1	13	0	16	13	18	0	16	0	15	16	0
210	16	17	11	16	1	11	2	16	0	16	2	0	16
212	0	1	13	0	16	13	18	0	16	0	15	16	0
205	15	16	11	15	1	11	4	15	2	15	0	2	15
301	16	17	11	16	1	11	2	16	0	16	2	0	16
302	0	1	13	0	16	13	18	0	16	0	15	16	0

Table B.4 *Nucleotide (a) and amino acid (b)substitution tables for patient 4 (C1).*

*Boxed regions represent intraisolate sequence diversity. * T = no. of days post-OLT; R = ratio of quasispecies; Q = designation of representative quasispecies.*

(a)

T*:			-11				8		86	218			329			
R*:	7:	2:	1:	1:	1:	1	10:	1	(8)	9:	1:	1	7:	2:	2:	1
Q*:	002	004	010	012	013	014	115	116	605	702	704	707	801	804	809	808
002	0	3	5	4	6	1	5	4	4	4	5	5	4	3	5	5
004	3	0	6	5	7	4	6	5	5	5	6	6	5	4	4	6
010	5	6	0	1	3	6	2	1	1	1	2	2	1	2	2	2
012	4	5	1	0	2	5	1	0	0	0	1	1	0	1	1	1
013	6	7	3	2	0	7	1	2	2	2	3	3	2	3	3	3
014	1	4	6	5	7	0	6	5	5	5	6	6	5	4	6	4
115	5	6	2	1	1	6	0	1	1	1	2	2	1	2	2	2
116	4	5	1	0	2	5	1	0	0	0	1	1	0	1	1	1
605	4	5	1	0	2	5	1	0	0	0	1	1	0	1	1	1
702	4	5	1	0	2	5	1	0	0	0	1	1	0	1	1	1
704	5	6	2	1	3	6	2	1	1	1	0	2	1	2	2	2
707	5	6	2	1	3	6	2	1	1	1	2	0	1	2	2	2
801	4	5	1	0	2	5	1	0	0	0	1	1	0	1	1	1
804	3	4	2	1	3	4	2	1	1	1	2	2	1	0	2	2
809	5	4	2	1	3	6	2	1	1	1	2	2	1	2	0	2
808	5	6	2	1	3	4	2	1	1	1	2	2	1	2	2	0

(b)

T*			-11				8		86	218			329			
R*	7:	2:	1:	1:	1:	1	10:	1	(8)	9:	1:	1	7:	2:	2:	1
Q*	002	004	010	012	013	014	115	116	605	702	704	707	801	804	809	808
002	0	0	1	1	1	0	1	1	1	1	1	2	1	1	1	1
004	0	0	1	1	1	0	1	1	1	1	1	2	1	1	1	1
010	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
012	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
013	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
014	0	0	1	1	1	0	1	1	1	1	1	2	1	1	1	1
115	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
116	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
605	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
702	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
704	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
707	2	2	1	1	1	2	1	1	1	1	1	0	1	1	1	1
801	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
804	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
809	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0
808	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0

Table B.5 *Nucleotide (a) and amino acid (b)substitution tables for patient 5 (L2)*

*Boxed regions represent intraisolate sequence diversity. * T = no. of days post-OLT; R = ratio of quasispecies; Q = designation of representative quasispecies.*

(a)

T*:			31				97			180		334		435			
R*:	8:	2:	1:	1:	1	6:	5:	1	10:	3	13:	1	11:	1:	1:	1
Q*:	A05	A04	A07	A13	A14	C02	C03	C16	F02	F05	Y03	Y10	102	109	115	118
A05	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
A04	1	0	9	2	14	1	3	1	1	2	1	2	2	3	3	2
A07	10	9	0	11	5	10	11	10	10	11	10	11	11	10	10	11
A13	1	2	11	0	16	1	5	3	1	2	1	2	2	3	3	2
A14	15	14	5	16	0	15	15	15	15	16	15	16	16	16	15	16
C02	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
C03	4	3	11	5	15	4	0	4	4	5	4	4	5	6	6	5
C16	2	1	10	3	15	2	4	0	2	3	2	3	3	4	4	3
F02	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
F05	1	2	11	2	16	1	5	3	1	0	1	2	0	1	1	2
Y03	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
Y10	1	2	11	2	16	1	4	3	1	2	1	0	2	3	3	2
102	1	2	11	2	16	1	5	3	1	0	1	2	0	1	1	2
109	2	3	10	3	16	2	6	4	2	1	2	3	1	0	2	3
115	2	3	10	3	15	2	6	4	2	1	2	3	1	2	0	3
118	1	2	11	2	16	1	5	3	1	2	1	2	2	3	3	0

(b)

T*:			31				97			180		334		435			
R*:	8:	2:	1:	1:	1	6:	5:	1	10:	3	13:	1	11:	1:	1:	1
Q*:	A05	A04	A07	A13	A14	C02	C03	C16	F02	F05	Y03	Y10	102	109	115	118
A05	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
A04	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
A07	10	9	0	11	5	10	11	10	10	11	10	11	11	10	10	11
A13	1	2	11	0	16	1	5	3	1	2	1	2	2	3	3	2
A14	15	14	5	16	0	15	15	15	15	16	15	16	16	16	15	16
C02	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
C03	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
C16	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
F02	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
F05	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
Y03	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
Y10	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
102	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
109	1	1	6	2	8	1	1	1	1	1	1	1	1	0	1	1
115	1	1	7	2	8	1	1	1	1	1	1	1	1	1	0	1
118	1	1	7	1	9	1	1	1	1	1	1	1	1	1	1	0

Table B.6 *Nucleotide (a) and amino acid (b)substitution tables for patient 6 (L2)*

Boxed regions represent intraisolate sequence diversity. * *T = no. of days post-OLT; R = ratio of quasispecies; Q = designation of representative quasispecies.*

(a)

T*:	31					97			180		334		435			
R*:	8:	2:	1:	1:	1	6:	5:	1	10:	3	13:	1	11:	1:	1:	1
Q*:	A05	A04	A07	A13	A14	C02	C03	C16	F02	F05	Y03	Y10	102	109	115	118
A05	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
A04	1	0	9	2	14	1	3	1	1	2	1	2	2	3	3	2
A07	10	9	0	11	5	10	11	10	10	11	10	11	11	10	10	11
A13	1	2	11	0	16	1	5	3	1	2	1	2	2	3	3	2
A14	15	14	5	16	0	15	15	15	15	16	15	16	16	16	15	16
C02	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
C03	4	3	11	5	15	4	0	4	4	5	4	4	5	6	6	5
C16	2	1	10	3	15	2	4	0	2	3	2	3	3	4	4	3
F02	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
F05	1	2	11	2	16	1	5	3	1	0	1	2	0	1	1	2
Y03	0	1	10	1	15	0	4	2	0	1	0	1	1	2	2	1
Y10	1	2	11	2	16	1	4	3	1	2	1	0	2	3	3	2
102	1	2	11	2	16	1	5	3	1	0	1	2	0	1	1	2
109	2	3	10	3	16	2	6	4	2	1	2	3	1	0	2	3
115	2	3	10	3	15	2	6	4	2	1	2	3	1	2	0	3
118	1	2	11	2	16	1	5	3	1	2	1	2	2	3	3	0

(b)

T*:	31					97			180		334		435			
R*:	8:	2:	1:	1:	1	6:	5:	1	10:	3	13:	1	11:	1:	1:	1
Q*:	A05	A04	A07	A13	A14	C02	C03	C16	F02	F05	Y03	Y10	102	109	115	118
A05	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
A04	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
A07	10	9	0	11	5	10	11	10	10	11	10	11	11	10	10	11
A13	1	2	11	0	16	1	5	3	1	2	1	2	2	3	3	2
A14	15	14	5	16	0	15	15	15	15	16	15	16	16	16	15	16
C02	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
C03	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
C16	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
F02	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
F05	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
Y03	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
Y10	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
102	0	0	6	1	8	0	0	0	0	0	0	0	0	1	1	1
109	1	1	6	2	8	1	1	1	1	1	1	1	1	0	1	1
115	1	1	7	2	8	1	1	1	1	1	1	1	1	1	0	1
118	1	1	7	1	9	1	1	1	1	1	1	1	1	1	1	0

Table B.7 *Nucleotide (a) and amino acid (b)substitution tables for patient 7 (L3)*

*Boxed regions represent intraisolate sequence diversity. * T = no. of days post-OLT; R = ratio of quasispecies; Q = designation of representative quasispecies.*

(a)

T*	0		47	
R*	9:	1	11:	1
Q*	001	022	101	107
001	0	3	3	3
022	3	0	0	4
101	3	0	0	4
107	3	3	4	0

(b)

T*	0		47	
R*	9:	1	11:	1
Q*	001	022	101	107
001	0	2	2	3
022	2	0	0	4
101	2	0	0	4
107	3	4	4	0

Appendix C:
Raw data from electronic cell counting device, indicating purity of leukocyte sub-populations prepared from two HCV-infected patients by the method of section 2.3.2

Patient A1: Pre-OLT sample

```
                RUN RESULTS
                WHOLE BLOOD

DATE: 02/21/94
TIME: 12:40

PATIENT ID#: LA1B2 (Lymphocytes)
SEQUENCE NO.   187
CROWN NO.
TUBE NO.

                                    M F
                                    -----
WBC      0.4 10^9/L
RBC     0.00 10^12/L
HGB      0.0 G/DL
HCT          %
MCV         FL
MCH         PG
MCHC        G/DL
RDW          %
PLT        5 10^9/L

MPV         FL

%LYMPH  33.6 %                      R
%MID     4.9 %                      R
%GRAN   11.5 %                      R
LYMPH    0.3 10^9/L                 R
MID      0.0 10^9/L                 R
GRAN     0.1 10^9/L                 R

REVIEW DIFF
```

```
                RUN RESULTS
                WHOLE BLOOD

DATE: 02/21/94
TIME: 12:46

PATIENT ID#: OA1B2 (Polymorphs)
SEQUENCE NO.   188
CROWN NO.
TUBE NO.

                                    M F
                                    -----
WBC      5.8 10^9/L
RBC     0.01 10^12/L
HGB      0.0 G/DL
HCT          %
MCV         FL
MCH         PG
MCHC        G/DL
RDW          %
PLT        3 10^9/L

MPV         FL

%LYMPH   4.2 %                      R
%MID     3.8 %                      R
%GRAN   92.0 %                      R
LYMPH    0.2 10^9/L                 R
MID      0.2 10^9/L                 R
GRAN     5.4 10^9/L                 R

LOW LYMPH MODE
```

Patient F1: Post-OLT sample.

```
                RUN RESULTS
                WHOLE BLOOD

DATE: 02/21/94
TIME: 12:49

PATIENT ID#: LF1A2 (Lymphocytes)
SEQUENCE NO.   189
CROWN NO.
TUBE NO.

                                    M F
                                    -----
WBC      0.9 10^9/L
RBC     0.00 10^12/L
HGB      0.0 G/DL
HCT          %
MCV         FL
MCH         PG
MCHC        G/DL
RDW          %
PLT        7 10^9/L

MPV         FL

%LYMPH  75.8 %                      R
%MID     9.9 %                      R
%GRAN   14.3 %                      R
LYMPH    0.7 10^9/L                 R
MID      0.1 10^9/L                 R
```

```
                RUN RESULTS
                WHOLE BLOOD

DATE: 02/21/94
TIME: 12:51

PATIENT ID#: OF1A2 (Polymorphs)
SEQUENCE NO.   190
CROWN NO.
TUBE NO.

                                    M F
                                    -----
WBC      4.1 10^9/L
RBC     0.01 10^12/L
HGB      0.0 G/DL
HCT          %
MCV         FL
MCH         PG
MCHC        G/DL
RDW          %
PLT        2 10^9/L

MPV         FL

%LYMPH   5.7 %                      R
%MID     4.9 %                      R
%GRAN   89.4 %                      R
LYMPH    0.2 10^9/L                 R
MID      0.2 10^9/L                 R
```

APPENDIX D

HCV viraemia levels in HCV-infected asymptomatic subjects and OLT recipients

Patient	HCV genotype	Viraemia level (genome equivalents/ml)				
Non-recipients of OLT:						
1 (Wad)	1a	0	267			
		ND	$7.0x10^5$			
2 (Dix)	2b	0	457	719		
		ND	$8.0x10^6$	$6.0x10^6$		
3 (Fra)	1b	0	24	87		
		$7.0x10^6$	$1.5x10^7$	$5.0x10^6$		
OLT recipients:						
4 (C1)	1b	-11[1]	8	86	218	329
		$2.5x10^4$	$5.0x10^5$	$7.0x10^6$	$5.5x10^6$	$7.0x10^6$
5 (L2)	1b	31	97	180	334	465
		$1.0x10^5$	$5.2x10^6$	ND[2]	ND	ND
6 (L1)	1b	-33	88			
		$5.0x10^6$	$7.0x10^8$			
7 (L3)	1a	-33	47			
		$<1x10^3$	$4.5x10^5$			

[1]. Time from OLT in days (- indicates prior to OLT). [2]. Not done.

Quantitation was carried out by Dr. Dazhuang Shang using a chemoluminescent detection method (Petrik *et al.* 1996).